MEMORANDUM 33

SENSITIVITY OF THE KALMAN FILTER WITH RESPECT TO PARAMETER VARIATIONS

Distribution:

Gardiner
Dressler
Meier
Ho
Kashiwagi
Keckler
Korsak
Lombard
Luenberger
Ratner

Salmon
Wong
Eikelman
Fraser
Yabroff
Albritton
Belote
Roscher
Weinstein
Branch

Davidson, NXPO
Montgomery, NXPO
Smith, NXPO
Wells, NXPO
Young, NXPO

Prepared by:

B. L. Ho
Research Engineer
Information and Control Laboratory

March 1968

Project 5188-305

SRI-H-8-984
R3

88

AD NUMBER	DATE	DDC ACCESSION NOTICE

1. REPORT IDENTIFYING INFORMATION

A. ORIGINATING AGENCY
Stanford Research Institute

B. REPORT TITLE AND/OR NUMBER "Sensitivity of Kalman Filter with Respect to Parameter Variations"

C. MONITOR REPORT NUMBER
Project 5188-305

D. PREPARED UNDER CONTRACT NUMBER(S)
DA-01-021-AMC-90006(Y)

2. DISTRIBUTION STATEMENT Statement A: Approved for Public Release; Distribution Unlimited

DDC FORM 50 AUG 65

DDC ACCESSION NOTICE

— INSTRUCTIONS —

Requester:

1. Put YOUR mailing address on reverse of form.
2. Complete items 1 and 2.
3. Attach form to reports mailed to DDC.
4. Use UNCLASSIFIED information only.

DDC:

1. Assign AD number.
2. Return to requester.

Best
Available
Copy

CONTENTS

ILLUSTRATIONS

I INTRODUCTION

The performance of a Kalman filter (recursive filter) depends on a number of parameters, such as measurement accuracy, a priori statistics of the initial state, model accuracy, sampling period, and methods of computing filter gains. Some of these are under the jurisdiction of the designer while others are specified, either as fixed values or as a range of values. For the designer to make an appropriate choice of the parameters, he needs to know the sensitivity of the filter performance with respect to the stated parameters.

The sensitivity information will answer, as examples, the following questions frequently encountered in the design of a Kalman filter.

(1) The noise statistics used in the filter need not be the exact statistics. The question is: How much performance improvement may be obtained if the statistics are more accurate?

(2) The iteration rate in computation (sampling rate) need not be identical to the data rate. The questions are: In the interest of computational requirements, how slow can the iteration rate be? How shall the data between the iterations be treated?

(3) The dynamics equations in the reference model of the filter need only be an approximation of the actual equations. The question is: How much approximation is tolerable?

(4) A significant part of filter computation is for the optimum filter gain, yet it has been shown[1*] that suboptimum gain requiring fewer computations are frequently just as good. The question is: What is the effect of suboptimal gains?

---

*References are listed at the end of the text.

(5) Frequently, a certain type of observations is expensive to make; therefore, it should be used with discretion. The question is: When should we use this observation?

(6) In implementing a real-time Kalman filter on a computer with limited word length, computation noises are introduced. The questions are: What is the acceptable word length? Should the filter gain be adjusted to account for the computation noise?

The sensitivity of Kalman filter may also be applied to analyze existing filters that are not of the Kalman type. Specifically, the difference in the estimation errors for the two filters may be computed from the sensitivities. Since Kalman filter is known to be optimum, we can then judge how close to optimum is the existing filter's performance.

Derivations of techniques and equations for the stated sensitivity questions are the main concern of this memorandum. Special emphasis is placed on the sampling-period sensitivities. Examples are given mainly to illustrate the applications of these results. The techniques can be incorporated readily into the existing computer programs.

The organization of this memorandum is escribed in the following.

The error covariances computed by the well-known covariance-matrix equations are used as the basis of filter performance. The validity of this approach is discussed in Sec. II-A and in Appendix A. In this regard, one should make a clear distinction between the "actual" covariance equations, which give the statistical descriptions of the actual errors, and the computed covariance equations, whose main purpose is to obtain the filter gains.

In gain computations, one obtains certain matrices, which are loosely called the covariance matrices. These may not represent the actual covariances of the estimation error because erroneous noise and dynamics parameter values may be used in the computation. This may be done unintentionally because knowledge of the process is imperfect, or intentionally to reduce the computation load. In this memo, all

covariances and equations for their computations shall refer to the actual ones, unless specific reference is made to the contrary.

The exact changes in the error covariances due to parameter variations are given in Sec. II-B. Of special interest to the designer is the question of gain variation. The designer is likely to have more control of the filter gain parameter than the other parameters, such as the measurement noise. For example, when the parameters change, he may either readjust the filter gains to optimal, or not change the gain at all. This and other special cases of filter gain can usually be studied better by using the specialized equations given in Sec. II-B.

The stated error-covariance-variation problem may be considered as a special case of the combined optimal control and estimation sensitivity, a continuous-time treatment of which is available.[2] On the other hand, most results available in the literature[3-7] may be considered a special case of this memorandum--parameters fixed and filter gains varied from optimal to arbitrary.

If estimation error at a specific time, say the terminal time, is of interest, the adjoint matrix technique may be used to facilitate the analysis. The computational advantage will be especially significant whenever an extensive analysis is to be made. This is also discussed in Sec. II-B with details in Appendix C.

Sections II-C, -D, and -E are motivated by the sampling-period variations. Two characteristics of the sampling period call for the special analysis of these paragraphs. First, it is a scalar parameter so that the sensitivity matrices or sensitivity indices may be defined as in Secs. II-D and -E. This is not possible, for example, for noise covariance-matrix variations. Secondly, for a fixed number of stages, changing the sampling period changes the overall duration of the process. Therefore, if the overall time duration were to remain fixed, reducing the sampling period would mean more sample points are available for filtering. In evaluating performance changes, this may have to be taken into account.

In Sec. II-F, we give the second-order effects of filter gain variation. This is necessary because, if the nominal filter gain is optimum, the first-order effects are zero.

Section III specializes the previous results to the sampling-period variations. In Sec. III-A, it is suggested that the sensitivity matrices, as defined in Sec. II-C, may be used in a gradient procedure for the design of sampling period. The sensitivity matrix is computed in the forward direction, and therefore may be most easily incorporated into existing programs. The restriction of this technique is that the sampling-period variations must be a function of a scalar variable--for example, uniform sampling period.

Nonuniform (unequal) sampling periods may be beneficial if, because of computational limitations, the number of stages are an important consideration. This is discussed in Sec. III-B. Suppose one is interested in the error covariance at a certain time T. All filtering done prior to T contributes to the reduction in the error covariance at T, but the filtering done during certain segments of this time may contribute more to this reduction than for other time segments. The sensitivity index, derived from the adjoint matrices, will be a useful technique here. The adjoint matrix is computed backwards starting from T; it relates the covariance variations at T to covariance variations at any stage prior to T.

In Sec. III-C, we point out possible applications of the sensitivity techniques to real-time computation allocation in a multiple-threat estimation situation--many separate estimation tasks are being performed simultaneously (for example, the multiple-threat situation in the anti-missile missile system). At any instant of time, the effectiveness of the Kalman filter for different targets will vary. Using the sensitivity indices, we may formulate and optimize the computation allocation problem.

According to a recent study,[8] a sudden unpredictable change of plant noise (such as a missile maneuver) may be detected from the measurements. In the same study, the noise increase is compensated by changing the filter gain so that it is optimum for the increased noise. However, the

response time of the filter is seen to be limited by the sampling period, even with the reoptimized gains. Further improvement is expected to lie in the reduction of the sampling period. The amount of reduction may be estimated from the trade-off between plant noise and sampling period using the sensitivity information as discussed in Sec. III-D.

In Sec. III-E, an example is presented for a $1/s^2$ plant with position measurements in which the sensitivity techniques that are developed are applied to the problem of reducing the number of measurements to be processed.

The trade-off between various parameters, as we have already mentioned, may be based on sensitivity information. Section IV describes this approach. In Sec. IV, sensitivity equations are given for the $1/s^2$ plant mentioned earlier. The results given there can be used to provide insight into more complicated problems. Also, the one-stage reduction in position error is plotted for various parameters in normalized quantities.

II ANALYTICAL RESULTS

A. Covariance Equations as a Measure of Estimation Error

Let the actual system be

$$
\begin{aligned}
x_{k+1} &= f_k(x_k) + w_k \\
z_k &= h_k(x_k) + v_k \quad ,
\end{aligned}
\tag{1}
$$

where f_k and h_k are (in general) nonlinear functions, x_k is the state to be estimated, z_k is the measurement, w_k and v_k are respectively the plant and measurement noises with zero-mean Gaussian distribution; they are mutually uncorrelated as well as time-uncorrelated:

$$
\begin{aligned}
&E[w_k] = 0 \quad ; \quad E\left[w_i w_j^T\right] = 0 \text{ for } i \neq j \\
&E\left[w_k w_k^T\right] = Q_k \\
&E[v_k] = 0 \quad ; \quad E\left[v_i v_j^T\right] = 0 \text{ for } i \neq j \\
&E\left[v_k v_k^T\right] = R_k \\
&E\left[v_i w_j^T\right] = 0 \text{ for all } i, j \quad .
\end{aligned}
$$

The estimate of x_k, given measurements $z_0, z_1, \ldots, z_k$, is denoted by $\hat{x}_{k/k}$. The extended Kalman filter yields the following recursive estimation equations:[1,9]

$$
\begin{aligned}
\hat{x}_{0/-1} &= E[x_0] = \bar{x}_0 \\
\hat{x}_{k/k-1} &= f_{k-1}(\hat{x}_{k-1/k-1}) \\
\hat{z}_{k/k-1} &= h_k(\hat{x}_{k/k-1}) \\
\hat{x}_{k/k} &= \hat{x}_{k/k-1} + W_k(z_k - \hat{z}_{k/k-1}) \quad ,
\end{aligned}
\tag{2}
$$

where W_k is the filter gain computed from a set of covariance equations:

$$P_{k/k-1} = \Phi_{k-1} P_{k-1/k-1} \Phi_{k-1}^T + Q_{k-1}$$
$$P_{k/k} = (I - W_k H_k) P_{k/k-1} (I - W_k H_k)^T + W_k R_k W_k^T \quad , \tag{3}$$

where

$$\Phi_{k-1} = \left(f_{k-1}\right)_x = \partial f_{k-1}/\partial x$$
$$H_k = \left(h_k\right)_x = \partial h_k/\partial x,$$

and the initial value for the covariance is

$$P_{0/-1} = E\left[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T\right].$$

In Eq. (3), the filter gain W_k is arbitrary. It is common practice to use the following approximately optimum* gain:

$$W_k = P_{k/k-1} H_k^T \left(H_k P_{k/k-1} H_k^T + R_k\right)^{-1} \quad . \tag{4}$$

In this memorandum, the error covariance is taken to be a measure of the performance of the extended Kalman filter. The estimation error, denoted by $\tilde{x}_{k/k}$ or $\tilde{x}_{k/k-1}$, is defined as

$$\tilde{x}_{k/k} = x_k - \hat{x}_{k/k}$$
$$\tilde{x}_{k/k-1} = x_k - \hat{x}_{k/k-1} \quad . \tag{5}$$

Since they are random variables, descriptions of $\tilde{x}_{k/k}$ and $\tilde{x}_{k/k-1}$ require a description of their probability distributions.

---

*The gain in Eq. (4) yields a minimum variance estimator, or maximum likelihood estimator for the linearized system of Eq. (1) under the stated assumptions on the noises v and w.

When f_k and h_k are linear functions, it can be shown that $\tilde{x}_{k/k}$ is a Gaussian variable. Its expectation and covariance are, therefore, sufficient to describe its probability distribution. In fact,

$$\bar{\tilde{x}}_{k/k} = E[\tilde{x}_{k/k}] = 0 \tag{6}$$

$$E\left[(\tilde{x}_{k/k} - \bar{\tilde{x}}_{k/k})(\tilde{x}_{k/k} - \bar{\tilde{x}}_{k/k})^T\right] = P_{k/k} \quad , \tag{7}$$

where $P_{k/k}$ is given in Eq. (3).

When f_k and h_k are nonlinear functions, as generally is the case in practical systems, $\tilde{x}_{k/k}$ is no longer Gaussian. In such cases, it may be difficult to obtain a complete description of its probability distributions. However, the first two moments $E[\tilde{x}_{k/k}]$ and $E\left[(\tilde{x}_{k/k} - \bar{\tilde{x}}_{k/k})(\tilde{x}_{k/k} - \bar{\tilde{x}}_{k/k})^T\right]$ give a fairly good idea of the probability distribution. As shown in Appendix A, they are easily approximated by

$$\begin{aligned}
E[\tilde{x}_{k/k-1}] \triangleq \bar{\tilde{x}}_{k/k-1} &= \Phi_{k-1}\bar{\tilde{x}}_{k-1/k-1} \\
&\quad + 1/2\left(f_{k-1}\right)_{xx} \circ \left[P_{k-1/k-1} + \bar{\tilde{x}}_{k-1/k-1}\bar{\tilde{x}}^T_{k-1/k-1}\right] \\
&\quad + 1/2\ Q_{k-1} \\
E[\tilde{x}_{k/k}] \triangleq \bar{\tilde{x}}_{k/k} &= (I - W_k H_k)\,\bar{\tilde{x}}_{k/k-1} \\
&\quad - 1/2\ W_k \left(h_k\right)_{xx} \circ \left[P_{k/k-1} + \tilde{x}_{k/k-1}\bar{\tilde{x}}^T_{k/k-1}\right] \\
&\quad - 1/2\ W_k R_k \quad ,
\end{aligned} \tag{8}$$

where we have used the following notation: $f_{xx} \circ P$ is a vector, the i-th component of which is given by

$$(f_{xx} \circ P)^{(i)} \doteq \sum_{j,k} \frac{\partial^2 f^{(i)}}{\partial x^{(j)} \partial x^{(k)}} P^{(j,k)}$$

$$E\left[(\tilde{x}_{k/k-1} - \bar{\tilde{x}}_{k/k-1})(\tilde{x}_{k/k-1} - \bar{\tilde{x}}_{k/k-1})^T\right] \triangleq P_{k/k-1} = \Phi_{k-1} P_{k-1/k-1} \Phi_{k-1}^T + Q_{k-1}$$

(9)

$$E\left[(\tilde{x}_{k/k} - \bar{\tilde{x}}_{k/k})(\tilde{x}_{k/k} - \bar{\tilde{x}}_{k/k})^T\right] \triangleq P_{k/k}$$

$$= (I - W_k H_k) P_{k/k-1} (I - W_k H_k)^T + W_k R_k W_k^T \quad .$$

We note that the approximate covariance equations in the nonlinear case of Eq. (9) are identical to Eq. (3), but the errors are no longer zero mean, as can be seen from Eq. (8).

The expected values in Eq. (8) may be made zero by modifying the extended Kalman filter equation of Eq. (1) to those corresponding to Eq. (A.11) in Appendix A. Normally, Eq. (8) is not available in an extended Kalman filter program while the covariances of Eq. (9) are available. For these reasons, the variations in covariances due to parameter variations will be analyzed in this memorandum, using Eq. (9).

B. Variations of the Error Covariance Matrix

We assume the existence of a nominal Kalman filter. This means that we use a nominal W_k (not necessarily optimum) in the manner shown in Eq. (2) for a physical process with parameters $P_{0/-1}$, Φ_k, Q_k, R_k. The resultant filter-error covariance is $P_{k/k}$. Let the physical parameters as well as the filter gain be changed so that

$$\text{Initial error } P_{0/-1} \to P_{0/-1} + \Delta P_{0/-1}$$

$$\text{Transition matrix } \Phi_k \to \Phi_k + \Delta\Phi_k$$

$$\text{Plant noise } Q_k \to Q_k + \Delta Q_k \tag{10}$$

$$\text{Measurement noise } R_k \to R_k + \Delta R_k$$

$$\text{Filter gain } W_k \to W_k + \Delta W_k \quad .$$

$P_{k/k}$ will be changed to $P_{k/k} + \Delta P_{k/k}$. We shall develop the formulas for $\Delta P_{k/k}$ and for $tr(L_{T/T}P_{T/T})$, where T is some fixed time and $L_{T/T}$ some appropriate matrix.

The effect of the variation $H_k \to H_k + \Delta H_k$ may always be viewed as the variation $\Phi_k \to \Phi_k + \Delta\Phi_k$ in an equivalent problem--for example, if the state x_{k+1} is expanded to include an additional term y_k, Eq. (1) may now be rewritten as

$$\begin{bmatrix} y_k \\ x_{k+1} \end{bmatrix} = \begin{bmatrix} h_k(x_k) \\ f_k(x_k) \end{bmatrix} + \begin{bmatrix} 0 \\ w_k \end{bmatrix}$$

$$z_k = [I \quad 0] \begin{bmatrix} y_k \\ x_{k+1} \end{bmatrix} + v_k \tag{11}$$

or, with obvious substitution of symbols,

$$x_k^* = f_{k-1}^*(x_{k-1}^*) + w_{k-1}^* \; ;$$

$$z_k = H_k x_k^* + v_k \quad . \tag{12}$$

An alternate way is to transform the state space coordinates so that the measurement variation appears as a dynamics variation. Thus,

let T_k be the transformation such that the outputs (minus noises) are components of a new state x_k^*:

$$x_k^* \triangleq T_k x_k = \begin{bmatrix} h_k(x_k) \\ y_k \end{bmatrix} \quad . \tag{13}$$

We may write Eq. (1) in the following form:

$$\begin{aligned} T_{k+1} x_{k+1} &= T_{k+1} f_k \left\{ T_k^{-1} \begin{bmatrix} h_k(x_k) \\ y_k \end{bmatrix} \right\} + T_{k+1} w_k \\ z_k &= [I \quad 0] \begin{bmatrix} h_k(x_k) \\ y_k \end{bmatrix} + v_k \end{aligned} \tag{14}$$

or

$$\begin{aligned} x_{k+1}^* &= f_k^*\left(T_k^{-1} x_k^*\right) + w_k^* \\ z_k &= H_k^* \, x_k^* + v_k \quad . \end{aligned} \tag{15}$$

In this new form, H_k^* is constant. This verifies our assertion that ΔH_k may be viewed as a $\Delta\Phi$.

Forward Recursive Equations

$\Delta P_{k/k}$ may be obtained by a straightforward substitution of Eq. (10) into the recursive covariance equations, as shown in Appendix B. The results may best be organized according to the optimality of the filter gains before and after parameters variations. Let us use the following notation:

W_{k+1}^o = Optimal gain for $P_{k/k}$, Φ_k, Q_k, and R_{k+1}

ΔW_{k+1}^o = Change in gain so that the new gain is optimal for

$$P_{k/k} + \Delta P_{k/k}, \ \Phi_k + \Delta\Phi_k, \ Q_k + \Delta Q_k, \text{ and } R_{k+1} + \Delta R_{k+1} \quad . \tag{16}$$

The deviation of the error covariance from its nominal is given by the expression

$$\Delta P_{k/k} = (I - W_k H_k)\ \Delta P_{k/k-1}\ (I - W_k H_k)^T + W_k \Delta R_k W_k^T + A_k + B_k \quad , \qquad (17)$$

where

$$\begin{aligned} A_k = & \left[W_k\left(H_k P_{k/k-1} H_k^T + R_k\right) - P_{k/k-1} H_k^T\right] \Delta W_k^T \\ & + \Delta W_k \left[W_k\left(H_k P_{k/k-1} H_k^T + R_k\right) - P_{k/k-1} H_k^T\right]^T \end{aligned} \qquad (18)$$

if

$$\Delta W_k = 0 \text{ or } W_k = W_k^o, \text{ we have } A_k = 0 \quad ; \qquad (19)$$

$$\begin{aligned} B_k = & \left[W_k\left(H_k \Delta P_{k/k-1} H_k^T + \Delta R_k\right) - \Delta P_{k/k-1} H_k^T\right] \Delta W_k^T \\ & + \Delta W_k \left[W_k\left(H_k \Delta P_{k/k-1} H_k^T + \Delta P_k\right) - \Delta P_{k/k-1} H_k^T\right]^T \\ & + \Delta W_k \left[H_k (P_{k/k-1} + \Delta P_{k/k-1}) H_k^T + (R_k + \Delta R_k)\right] \Delta W_k^T \quad . \end{aligned} \qquad (20)$$

Nominal quantities are used for $P_{k/k-1}$, H_k, R_k, and W_k. If $\Delta W_k = \Delta W_k^o$, we have the following negative definite expression for $(A_k + B_k)^o$:

$$\begin{aligned} (A_k + B_k)^o = - & \left\{\left[W_k\left(H_k \Delta P_{k/k-1} H_k^T + \Delta R_k\right) - \Delta P_{k/k-1} H_k^T\right]\right. \\ & \left. + \left[W_k\left(H_k P_{k/k-1} H_k^T + R_k\right) - P_{k/k-1} H_k^T\right]\right\} \\ & \cdot \left[H_k (P_{k/k-1} + \Delta P_{k/k-1}) H_k^T + (R_k + \Delta R_k)\right]^{-1} \\ & \cdot \left\{\left[W_k\left(H_k \Delta P_{k/k-1} H_k^T + \Delta R_k\right) - \Delta P_{k/k-1} H_k^T\right]^T\right. \\ & \left. + \left[W_k\left(H_k P_{k/k-1} H_k^T + R_k\right) - P_{k/k-1} H_k^T\right]^T\right\} \quad . \end{aligned} \qquad (21)$$

If $W_k = W_k^o$, the resultant $(A_k + B_k)^o$ will be denoted by B_k^o. If

$$\Delta W_k = 0, \text{ we have } B_k = 0 \quad . \tag{22}$$

If $\Delta W_k \neq 0$, but $\Delta P_{k/k-1}$, ΔR_k are zero, the expression for B_k becomes positive definite:

$$B_k = \Delta W_k \left[H_k P_{k/k-1} H_k^T + R_k \right] \Delta W_k^T \quad . \tag{23}$$

This equation and Eq. (8) in fact show that

$$W_k = P_{k/k-1} H_k^T \left(H_k P_{k/k-1} H_k^T + R_k \right)^{-1} \tag{24}$$

is a minimum variance filter gain for $P_{k/k-1}$ and R_k: if gain is changed from the given expression, the error covariance would be changed by an amount

$$\Delta P_{k/k} = A_k + B_k = 0 + \Delta W_k \left[H_k P_{k/k-1} H_k^T + R_k \right] \Delta W_k^T \quad , \tag{25}$$

which is positive definite, indicating an increase in the error covariance. To complete the recursive equations, $\Delta P_{k+1/k}$ is

$$\begin{aligned} \Delta P_{k+1/k} = {} & \Phi_k \, \Delta P_{k/k} \, \Phi_k^T + \Delta Q_k \\ & + \Delta \Phi_k (P_{k/k} + \Delta P_{k/k}) \, \Phi_k^T + \Phi_k (P_{k/k} + \Delta P_{k/k}) \, \Delta \Phi_k^T \\ & + \Delta \Phi_k (P_{k/k} + \Delta P_{k/k}) \, \Delta \Phi_k^T \quad . \end{aligned} \tag{26}$$

If $\Delta \Phi_k = 0$, we have

$$\Delta P_{k+1/k} = \Phi_k \Delta P_{k/k} \Phi_k^T + \Delta Q_k \quad . \tag{27}$$

Of particular interest is a comparison between the cases with $\Delta W_k = 0$ and $\Delta W_k = \Delta W_k^o$, when $W = W^o$. ΔW_k is zero when the filter is not

readjusted after parameter variations; we then have, by Eqs. (17), (19), and (22),

$$\Delta P_{k/k} = (I - W_k H_k)\ \Delta P_{k/k-1}(I - W_k H_k)^T + W_k \Delta R_k W_k^T \quad . \tag{28}$$

If the gain is readjusted to optimal, i.e., $\Delta w_k = \Delta w_k^o$, we have

$$\Delta P^o_{k/k} = (I - W_k H_k)\ \Delta P_{k/k-1}(I - W_k H_k)^T + W_k \Delta R_k W_k^T + B_k^o \quad , \tag{29}$$

where B_k^o is a negative definite matrix [see Eq. (21)]:

$$- \left[-(I - W_k H_k)\Delta P_{k/k-1} H_k^T + W_k \Delta R_k\right] \left[H_k(P_{k/k-1} + \Delta P_{k/k-1})H_k^T + (R_k + \Delta R_k)\right]^{-1}$$
$$\cdot \left[-(I - W_k H_k)\Delta P_{k/k-1} H_k^T + W_k \Delta R_k\right]^T \quad .$$

Because

$$\Delta P_{k/k} - \Delta P^o_{k/k} = - B_k^o \quad , \tag{30}$$

B_k^o, therefore, represents the decrease in the error covariance by optimal readjustment of the gains according to the variations in Φ_k, P_k, Q_k, and R_{k+1} over the case when the variations are ignored. B_k^o is of second- and third-order of the variations [see Eq. (21)]. When the first-order term dominates in problems such as the sampling-period sensitivity the B_k^o term may safely be ignored. On the other hand, this term is of utmost importance in analyzing suboptimal gains, because the first-order term is zero. From the value of information viewpoint, B_k^o justifies the accuracies the designer has available or is requesting on Φ_k, P_k, Q_k, and R_{k+1}. It tells how much improvement in estimation error he would obtain if the accuracies were improved.

The variational equations up to this point have been exact and general. In what follows, we may sometimes use specialized and/or approximate versions of the variational equations for specific problems.

These problems are organized and discussed in Secs. II-G and III, but first we shall introduce the concept of the adjoint matrix for the Kalman filter.

The Error Transition Matrix and the Adjoint Matrix

Often one wishes to know how $P_{T/T}$ at a specific sample point T changes because of changes in Q, R, Φ, or W at some previous and possibly widely separated sample points. Using the error transition matrices and adjoint matrices discussed later in this section, it is possible to study these changes without having to solve repeatedly the full recursive Eqs. (17) and (26). However, for these techniques to be feasible, we have to use the first-order approximation (second-order approximation in suboptimal gain case) of the full recursive Eqs. (17) and (26).

Let us study the variations from k - 1 to kth sample points. The relevant variations are $\Delta\Phi_{k-1}$, ΔQ_{k-1}, ΔR_k, and ΔW_k. We also have $\Delta P_{k-1/k-1}$, which is caused by the previous parameter variations. We shall take $\Delta P_{k-1/k-1}$ to be first order in magnitude. This is not always so, as we shall see later on in suboptimal gain analysis that $\Delta P_{k-1/k-1}$ is second order in ΔW. Let us also define matrices $\Delta Z_{k/k-1}$ and $\Delta Z_{k/k}$ by the following equations, which may be regarded as the first-order approximation of the local (one-stage) covariance variations--by setting $\Delta P_{k-1/k-1} = 0$ in Eqs. (13) and (7) and keeping the first-order terms:

$$
\begin{aligned}
\Delta Z_{k/k-1} &= \Delta Q_{k-1} + \Delta\Phi_{k-1} P_{k-1/k-1} \Phi^T_{k-1} + \Phi_{k-1} P_{k-1/k-1} \Delta\Phi^T_{k-1} \\
\Delta Z_{k/k} &= (I - W_k H_k)\, \Delta Z_{k/k-1}\, (I - W_k H_k)^T + W_k \Delta R_k W^T_k + A_k \qquad (31) \\
\Delta Z_{0/-1} &\triangleq \Delta P_{0/-1} \quad ,
\end{aligned}
$$

where A_k is given in Eq. (18), a first-order expression in ΔW. B_k does not appear because it is higher than first order. The only parameter variations that appear in the above equations are local--$\Delta\Phi_{k-1}$, ΔQ_{k-1}, ΔR_k, and ΔW_k. Furthermore, $\Delta Z_{k/k} = 0$ if these local parameters are unchanged.

The covariance variations [Eqs. (17) and (26)] are now:

$$\Delta P_{k/k} = (I - W_k H_k)\Phi_{k-1}\Delta P_{k-1/k-1}\Phi^T_{k-1}(I - W_k H_k)^T + \Delta Z_{k/k} \quad . \qquad (32)$$

Applying the last equation recursively from k = 0 to k = T, we have

$$\Delta P_{T/T} = \sum_j D_{T,j}\, \Delta Z_{j/j}\, D^T_{T,j} \quad , \qquad (33)$$

where j = those integers such that at least one of the variations $\Delta\Phi_{j-1}$, ΔQ_{j-1}, ΔR_j, and ΔW_j exist. $D_{T,j}$ is given by Eq. (34):

$$D_{T,j} \triangleq (I - W_T H_T)\Phi_{T-1}\,(I - W_{T-1}H_{T-1})\Phi_{T-2}\,\cdots\,(I - W_{j+1}H_{j+1})\Phi_j \quad . \qquad (34)$$

The $D_{T,j}$ will be called the <u>error transition matrices</u>. They give a direct connection between the local covariance variations $\Delta Z_{j/j}$ and $\Delta P_{T/T}$. They are computed backwards, using nominal quantities in time from T to j by the following recursive equation:

$$D_{T,j} = D_{T,j+1}\,(I - W_{j+1}H_{j+1})\Phi_j$$
$$D_{T,T} = I, \text{ the unit matrix} \quad . \qquad (35)$$

In particular problems, it may be more convenient to use multistage local variations. The filtering process is divided (see Fig. 1) into m time segments; let the dividing points be the set of sample points $(j_1, j_2, j_3, \ldots, j_m)$. We have

$$\Delta P_{T/T} = \sum_i D_{T,j_i}\, \Delta Z^{(j_{i-1})}_{j_i/j_i}\, D^T_{T,j_i} \quad , \qquad (36)$$

where $\Delta Z^{(j_{i-1})}_{j_i/j_i}$ denotes a multistage local variation for the j_{i-1} to j_ith segment, obtained by solving the following recursive equations [Eq. (37)] $(j_i - j_{i-1})$ times:

$$\Delta Z_{k/k-1}^{(j_{i-1})} = \Phi_{k-1} \Delta Z_{k-1/k-1}^{(j_{i-1})} \Phi_{k-1}^T + \Delta Q_{k-1} + \Delta\Phi_{k-1} P_{k-1/k-1} \Phi_{k-1}^T + \Phi_{k-1} P_{k-1/k-1} (\Delta\Phi_{k-1})^T$$

$$\Delta Z_{k/k}^{(j_{i-1})} = (I - W_k H_k) \Delta Z_{k/k-1}^{(j_{i-1})} (I - W_k H_k)^T + W_k \Delta R_k W_k^T + A_k \quad , \tag{37}$$

where A_k is defined in Eq. (18), and

$$k = j_{i-1} + 1, \; j_{i-1} + 2, \; \ldots, \; j_i$$

with

$$\Delta Z_{j_{i-1}/j_{i-1}}^{(j_{i-1})} = 0 \quad .$$

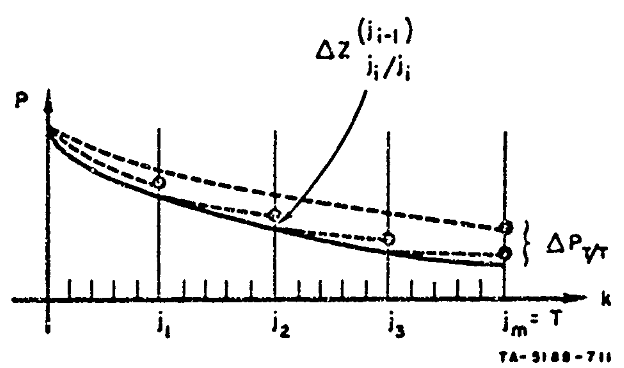

FIG. 1 MULTISTAGE LOCAL VARIATIONS

The discussed error transition matrices are valuable in filter synthesis. For example, in order to investigate the effect of m difference ΔR_0 on $\Delta P_{T/T}$, according to Eqs. (33) and (31), one only has to investigate the expression $D_{T,0} W_0 R_0 W_0^T D_{0,T}^T$ m times--a great time saving from having to solve the recursive equations [Eqs. (17) and (26)] m times from k = 0 to T.

Often one is interested in knowing the following scalar function, J, of $\Delta P_{T/T}$ rather than the whole $\Delta P_{T/T}$ matrix itself:

$$J = \mathrm{tr}\ (L_{T/T} P_{T/T}) \tag{38}$$

$$\Delta J = \mathrm{tr}\ (L_{T/T} \Delta P_{T/T}) \quad , \tag{39}$$

where tr denotes the matrix trace operation.

As an example, the so-called root-sum-squared (rss) position error may be written in the above form for J. Thus, the state of a trajectory in three-dimensional motion may be

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} ,$$

where x_1, x_2, x_3 are position and x_4, x_5, x_6 are velocity coordinates. The error in x is δx:

$$\delta x = \begin{bmatrix} \delta x_1 \\ \delta x_2 \\ \delta x_3 \\ \delta x_4 \\ \delta x_5 \\ \delta x_6 \end{bmatrix} .$$

$P_{T/T}$ is the covariance matrix of δx,

$$P_{T/T} = E[(\delta x - \overline{\delta x})(\delta x - \overline{\delta x})^T] \quad .$$

The rss position error is, by the definitions of rss and the trace operator,

$$(rss) = tr(L_{T/T} P_{T/T})$$

with

$$L_{T/T} = \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & 0 & \\ & & & & & 0 \end{bmatrix} \quad . \tag{40}$$

The variation in rss position error is now expressible by

$$\Delta J = \Delta(rss) = tr(L_{T/T} \Delta P_{T/T}) \quad .$$

$L_{T/T}$ is the adjoint matrix at T. Its exact form depends on the problem at hand. As another example, if we are only interested in the covariance of δx_1, $L_{T/T}$ should then be

$$L_{T/T} = \begin{bmatrix} 1 & & & & & \\ & 0 & & & & \\ & & 0 & & & \\ & & & 0 & & \\ & & & & 0 & \\ & & & & & 0 \end{bmatrix} \quad .$$

The equation for computing ΔJ is obtained by combining Eqs. (39) and (33):

$$\Delta J = \operatorname{tr}\left(L_{T/T}\sum_j D_{T,j}\Delta Z_{j/j}D_{T,j}^T\right)$$

$$= \sum_j \operatorname{tr}\left(L_{T/T}D_{T,j}\Delta Z_{j/j}D_{T,j}^T\right)$$

$$= \sum_j \operatorname{tr}\left(D_{T/j}^T L_{T/T}D_{T,j}\Delta Z_{j/j}\right)$$

$$\Delta J = \sum_j \operatorname{tr}(L_{j/j}\Delta Z_{j/j}) \tag{41}$$

$$L_{j/j} \triangleq D_{T,j}^T L_{T/T} D_{T,j} \quad . \tag{42}$$

The $L_{j/j}$ are called the adjoint matrices; $L_{j/j}$ may be computed [see Eq. (35)] by the following backwards recursive equation:

$$L_{k-1/k-1} = \Phi_{k-1}^T(I - W_kH_k)^T L_{k/k}(I - W_kH_k)\Phi_{k-1} \quad , \tag{43}$$

where $L_{T/T}$ is to be appropriately defined for the particular problem.

To summarize the above technique known as the adjoint-matrix technique, we first use Eqs. (42) and (43) to compute the adjoint matrices $L_{k/k}$, $k = T, T-1, T-2, \ldots$, using the nominal quantities W_k, H_k, and Φ_{k-1}. J may now be written easily using $L_{k/k}$ and local variations $\Delta Z_{j/j}$ of Eq. (31), thus

$$\operatorname{tr}(L_{T/T}\Delta P_{T/T}) = \Delta J = \sum_j \operatorname{tr}(L_{j/j}\Delta Z_{j/j}) \quad ,$$

where j are those samples points with nonzero $\Delta Z_{j/j}$.

$L_{j/j}$ and $D_{T,j}$ enable one to relate directly local covariance variations $\Delta Z_{j/j}$, with terminal-error variation. This provides a systematic and straightforward technique for such sensitivity and trade-off analysis as

(1) The necessary duration of the filtering process

(2) Seeking least sensitive spots on the filtering time axis for sampling rate reduction

(3) Redistributing the sampling points for improved terminal accuracy

(4) Optimizing the usage of an expensive observation--the seeking of sample points when the observation causes the largest terminal error reduction

(5) Optimum radar resource allocation

(6) Trade-offs between various error sources--measurement, modeling, and sampling rate.

These topics will be discussed in more detail in later sections, and in examples of Sec. III.

C. The Sensitivity Matrix--Variations Due to Scalar Parameters

Up to now, we have considered the independent variations to be in the general matrix form--$\Delta\Phi_{k-1}$, ΔQ_{k-1}, ΔW_k, ΔR_k. Here we introduce another level of parametrization; we let $\Delta\Phi_{k-1}$, ΔQ_{k-1}, ΔW_k, and ΔR_k be functions of a scalar variable denoted by α. For example, the sampling period (time between two consecutive sample points) in a filter with uniform sampling rate is such a parameter.

Let each element in the matrices $\Delta\Phi_{k-1}$, ΔQ_{k-1}, ΔW_k, and ΔR_k be functions of a scalar variable α. Each element of $P_{k/k}$ is, therefore, a function of α. This leads to the following definition of a sensitivity matrix:

$$\frac{\partial P_{k/k}}{\partial\alpha} \triangleq \left[\frac{\partial P_{k/k}^{(ij)}(\alpha)}{\partial\alpha}\right] = \lim_{\Delta\alpha\to 0} \frac{1}{\Delta\alpha}\,[P(\alpha + \Delta\alpha) - P(\alpha)] = \lim_{\Delta\alpha\to 0} \frac{1}{\Delta\alpha}\,\Delta P_{k/k}(\alpha) \ . \quad (44)$$

$(\partial P_{k/k}/\partial\alpha)$ propagates in time according to a set of recursive equations, which will be shown later. Once $(\partial P_{k/k}/\partial\alpha)$ is obtained, $\Delta P_{k/k}$ may be easily computed to first order as

$$\Delta P_{k/k}(\alpha) \cong (\partial P_{k/k}/\partial\alpha)\ \Delta\alpha \quad . \tag{45}$$

The recursive equations for the sensitivity matrix may be derived by taking limits [as indicated in Eq. (44)] of Eqs. (17) and (26):

$$\frac{\partial P_{k/k}}{\partial\alpha} = (I - W_k H_k)\ \frac{\partial P_{k/k-1}}{\partial\alpha}\ (I - W_k H_k)^T + W_k\ \frac{\partial R_k}{\partial\alpha}\ W_k^T + A_k' \tag{46}$$

and

$$\frac{\partial P_{k+1/k}}{\partial\alpha} = \Phi_k\ \frac{\partial P_{k/k}}{\partial\alpha}\ \Phi_k^T + \frac{\partial Q_k}{\partial\alpha} + \frac{\partial\Phi_k}{\partial\alpha}\ P_{k/k}\Phi_k + \Phi_k^T P_{k/k}\left(\frac{\partial\Phi_k}{\partial\alpha}\right)^T \tag{47}$$

$$\begin{aligned} A_k' = {} & \left[W_k\left(H_k P_{k/k-1} H_k^T + R_k\right) - P_{k/k-1}H_k^T\right]\left(\frac{\partial W_k}{\partial\alpha}\right)^T \\ & + \left(\frac{\partial W_k}{\partial\alpha}\right)\left[W_k\left(H_k P_{k/k-1} H_k^T + R_k\right) - P_{k/k-1}H_k^T\right]^T \quad , \end{aligned} \tag{48}$$

where all second- and third-order terms vanish because of the limiting process. These equations are seen to be similar to the covariance equations, Eq. (3). This indicates that in a simulation of the Kalman filter, the sensitivity-matrix computation may be added with a minimum amount of effort.

The scalar parameters we shall be concerned with are:

(1) Sampling period τ: In general, Φ_k, Q_k, and R_k are functions of τ. Φ_k, the transition matrix for the system dynamics, is a time integral depending on τ. Q_k generally increases with τ because a longer sampling period allows the dynamics error to build up between sampling points. The measurement error R_k changes with τ in a manner depending on the data presmoothing used with the filter. If no presmoothing is used, then R_k does not change with τ. To summarize:

$\partial\Phi_k/\partial\tau \neq 0$, $\partial Q_k/\partial\tau \neq 0$, and $\partial R_k/\partial\tau$ may or may not equal zero.

(2) Dynamic error parameter, d: The dynamics error is commonly modeled as plant noise. Q_k is, therefore, a function of d, while Φ and R are not:

$$\partial Q_k/\partial d \neq 0,\ \partial R_k/\partial d = 0,\ \partial\Phi_k/\partial d = 0 \quad .$$

(3) Measurement noise parameter, r: $\partial R_k/\partial r \neq 0$, but $\partial\Phi_k/\partial r = 0$, and $\partial Q_k/\partial r = 0$.

To compute the sensitivity matrices of these parameters, we use Eqs. (46) and (47) with $\alpha = \tau$, d, or r.

D. Special Considerations in the Sampling-Period Sensitivity Matrix--the Time-Based Sensitivity Matrix

The sensitivity $\partial P_{k/k}/\partial\tau$ discussed previously is evaluated with k fixed. The derivative is therefore evaluated along the slanted dotted line in the P vs. t diagram of Fig. 2. But we are often interested

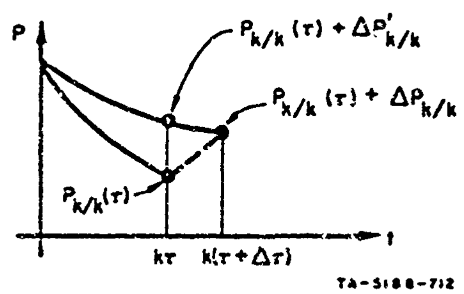

FIG. 2 TIME SHIFT IN SAMPLING-PERIOD VARIATION

in a sensitivity that is time-fixed; for example, consider the case in which the time duration of the filtering process is to remain unchanged. Then, in Fig. 2, we should consider a derivative evaluated along the vertical direction by letting $P_{k/k}(\tau) + \Delta P'_{k/k}$ approach $P_{k/k}(\tau)$. In this section, we shall devise techniques to evaluate the time-fixed sensitivity.

In Fig. 3 we depict the effect of a small change $\delta\tau$ in the sampling period τ for k iterations. $P_{k/k}(\tau)$, denoting the error covariance at k-th sample point with sampling period τ, is changed to $P_{k/k}(\tau + \Delta\tau) \triangleq P_{k/k}(\tau) + \Delta P_{k/k}$, and $P_{k-1/k-1}(\tau)$ to $P_{k-1/k-1}(\tau + \Delta\tau)$.

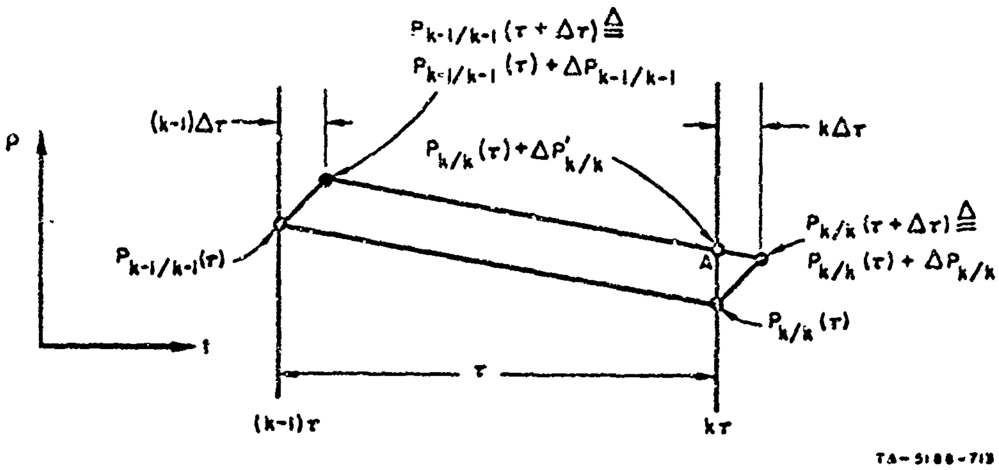

FIG. 3 INTERPOLATION BETWEEN SAMPLE POINTS

While, strictly speaking, the error covariances are not defined between the sample points k − 1 and k, we shall interpolate using $P_{k-1/k-1}(\tau + \Delta\tau)$ and $P_{k/k}(\tau + \Delta\tau)$ to assign a covariance matrix at the point marked A. This matrix is designated by $P'_{k/k}(\tau + \Delta\tau) = P_{k/k}(\tau) + \Delta P'_{k/k}$, and $\Delta P'_{k/k}$ is taken to be the fixed-time variation under the influence of $\Delta\tau$. Carrying out the interpolation, we have

$$P_{k/k}(\tau) + \Delta P'_{k/k} = [P_{k/k}(\tau) + \Delta P_{k/k}] + \frac{[P_{k-1/k-1}(\tau) + \Delta P_{k-1/k-1} - P_{k/k}(\tau) - \Delta P_{k/k}]}{(\tau + \Delta\tau)} k(\Delta\tau) .$$

Therefore,

$$\frac{\Delta P'_{k/k}}{\Delta\tau} = \frac{\Delta P_{k/k}}{\Delta\tau} + k\,\frac{P_{k-1/k-1}(\tau) - P_{k/k}(\tau)}{(\tau + \Delta\tau)} + k\,\frac{\Delta P_{k-1/k-1} - \Delta P_{k/k}}{(\tau + \Delta\tau)} \quad ; \qquad (49)$$

Letting $\Delta\tau \to 0$, and remembering that $\Delta P_{k-1/k-1} \to 0$, $\Delta P_{k/k} \to 0$ as $\Delta\tau \to 0$, so that the last term is zero as $\Delta\tau \to 0$, we have:

$$\frac{\partial P'_{k/k}}{\partial \tau} \triangleq \lim_{\Delta\tau \to 0} \frac{\Delta P'_{k/k}}{\Delta \tau} = \frac{\partial P_{k/k}}{\partial \tau} + \frac{k}{\tau} [P_{k-1/k-1}(\tau) - P_{k/k}(\tau)] \quad . \quad (50)$$

The time-fixed sampling-period sensitivity matrix $\partial P'_{k/k}/\partial\tau$ is therefore obtainable from the sampling-point-fixed sensitivity matrix $\partial P_{k/k}/\partial\tau$ by taking into account the slope of $P_{k/k}(\tau)$, as shown in Eq. (50).

E. The Sensitivity Index

If we limit our discussion to scalar parameter variations $\Delta\alpha$ and scalar filter performance criteria J, we are then interested in sensitivity indices of the type

$$S_\alpha \triangleq \frac{\partial J}{\partial \alpha} \quad , \quad (51)$$

where α = various scalar parameters, for example, those listed in the last part of Sec. II-C.

These sensitivity indices may be precomputed using the previous results on sensitivity and adjoint matrices; the S, being scalars, are stored more easily than the adjoint or sensitivity matrices. S are therefore seen to be a useful technique for real-time modifications of the filter. For example, in the antimissile defense system, we may let J be the covariance of the error in the position estimate at a certain range from the radar sites, where interceptor commitments have to be made. Let α be the iteration rate C. Now, assume a multiple-threat situation arises such that the design iteration rate C_o cannot be met for all incoming vehicles because of computation limitations. This calls for a real-time decision on how the iteration rate should be reduced, and an estimate of the consequences. These tasks are made easier if the sensitivity indices $S_c^{(j)}$ (j = 1, 2, ..., m for m different reentry vehicles) are available in real time, for then the increase in terminal position error for each vehicle is expressible as $\Delta J^{(j)} = S_c^{(j)} (\Delta C_j)$.

From criterion on $\Delta J^{(j)}$, because the terminal accuracies are the main concern in this example, we may compute ΔC_j from the stated equations.

To compute the sensitivity indices, we use the sensitivity matrices and adjoint-matrix technique discussed previously. This gives

$$S_\alpha = \frac{\partial J}{\partial \alpha} = \frac{\partial}{\partial \alpha} \operatorname{tr}(L_{T/T} P_{T/T}) = \operatorname{tr}\left(L_{T/T} \frac{\partial P_{T/T}}{\partial \alpha}\right)$$
$$= \sum_j \operatorname{tr}\left(L_{j/j} \frac{\partial Z_{j/j}}{\partial \alpha}\right) \quad , \tag{52}$$

where $(\partial Z_{k/k}/\partial\alpha)$ are the local one-stage sensitivity matrices obtained by the following limit:

$$\frac{\partial Z_{k/k}}{\partial \alpha} = \lim_{\Delta\alpha \to 0} \frac{\Delta Z_{k/k}}{\Delta \alpha} \quad . \tag{53}$$

Therefore, upon using Eq. (31), we have

$$\frac{\partial Z_{k/k-1}}{\partial \alpha} = \frac{\partial Q_{k-1}}{\partial \alpha} + \frac{\partial \Phi_{k-1}}{\partial \alpha} P_{k-1/k-1} \Phi_{k-1}^T + \Phi_{k-1} P_{k-1/k-1} \left(\frac{\partial \Phi_{k-1}}{\partial \alpha}\right)^T \tag{54}$$

$$\frac{\partial Z_{k/k}}{\partial \alpha} = (I - W_k H_k) \frac{\partial Z_{k/k-1}}{\partial \alpha} (I - W_k H_k)^T + W_k \frac{\partial R_k}{\partial \alpha} W_k^T + A_k'$$

$$A_k' = \left[W_k\left(H_k P_{k/k} H_k^T + R_k\right) - P_{k/k-1} H_k^T\right] \left(\frac{\partial W_k}{\partial \alpha}\right)^T$$
$$+ \left(\frac{\partial W}{\partial \alpha}\right) \left[W_k\left(H_k P_{k/k-1} H_k^T + R_k\right) - P_{k/k-1} H_k^T\right]^T \quad .$$

Note that the same equations may be obtained from the sensitivity equations [Eqs. (46) and (47)] upon setting

$$(\partial P_{k-1/k-1}/\partial\alpha) = 0 \quad .$$

If we are computing the sampling-period sensitivities for time-fixed variations, the slope correction shown in Eq. (50) must be used. Also, instead of using a one-stage local variation, $\Delta Z_{j/j}$, it may be

more convenient in particular problems to use a multistage local variation, $\Delta\Sigma_{i/i}^{(m)}$, of Eqs. (36) and (37).

F. Special Considerations in the Gain-Matrix Variation

If the nominal filter gain is optimum W_k^o, the sensitivity-matrix techniques such as those in Secs. II-C, -D, and -E, do not give usable answers, because in those techniques only first-order variations are considered, while the variation in performance about the optimum W_k^o is second order in ΔW_k. Here the analysis calls for a return to the exact equations of Sec. II-B for second-order variations. We shall consider only the special case that W_k varies from W_k^o to $W_k^o + \Delta W_k$; all other parameters (Φ_k, Q_k, R_{k+1}, $P_{0/1}$) are to remain fixed. Then, in the variational equations [Eqs. (17) and (36)] all first-order terms disappear. We have

$$\Delta P_{0/-1} = 0$$

$$\Delta P_{k/k-1} = \Phi_{k-1}\Delta P_{k-1/k-1}\Phi_{k-1}^T \tag{55}$$

$$\Delta P_{k/k} = \left(I - W_k^o H_k\right) \Delta P_{k/k-1}\left(I - W_k^o H_k\right)^T + \Delta W_k \left[H_k P_{k/k-1} H_k^T + R_k\right] \Delta W_k^T + \text{(higher-order terms)} \quad .$$

We have used the following observation in the last equation: $\Delta P_{k/k}$ and $\Delta P_{k+1/k}$ stays second order for $k = 0, 1, 2, \ldots$ as may be seen by tracing the recursive equation for $k = 0, 1, 2, \ldots$. Therefore, although the $\Delta P_{k/k-1} H_k^T \Delta W_k^T$ term in B_k, Eq. (20), appears to be of second order, it is really of third order.

Equation (55) is now approximated by neglecting the higher-order terms, obtaining

$$\Delta P_{k/k} = \left(I - W_k^o H_k\right) \Phi_{k-1}(\Delta P_{k-1/k-1}) \Phi_{k-1}^T \left(I - W_k^o H_k\right)^T + \Delta W_k\left(H_k P_{k/k-1} H_k^T + R_k\right) \Delta W_k^T \quad . \tag{56}$$

It is noted that, due to the separation of $\Delta P_{k-1/k-1}$ and ΔW_k terms, we may employ the concepts of the local covariance matrices, the error transition matrices, and the adjoint matrices--concepts that were developed originally for the first-order variation. The applicable equations are given below:

The local covariance variation is

$$\Delta Z_{k/k} = \Delta W_k \left(H_k P_{k/k-1} H_k^T + R_k \right) \Delta W_k^T \quad , \qquad k = 0, 1, 2, \ldots \tag{57}$$

so that

$$\Delta P_{k/k} = (I - W_k H_k)\ \Phi_{k-1} (\Delta P_{k-1/k-1})\ \Phi_{k-1}^T (I - W_k H_k)^T + \Delta Z_{k/k} \tag{58}$$

Terminal covariance variation is expressible as sum of local variations:

$$\Delta P_{T/T} = \sum_k D_{T,k} \Delta Z_{k,k} D_{T,k}^T \quad , \tag{59}$$

where $D_{T,k}$ is the error transition matrix by the recursive equation

$$D_{T,k} = D_{T,k+1} (I - W_{k+1} H_{k+1})\ \Phi_k \tag{60}$$

with

$$D_{T,T} = I \quad .$$

Terminal ΔJ is expressible as sum of local variations:

$$\Delta J = \mathrm{tr}(L_{T/T} \Delta P_{T/T}) = \sum_k \mathrm{tr}(L_{k/k} \Delta Z_{k/k}) \quad , \tag{61}$$

where $L_{k/k}$ is the adjoint matrix defined by the recursive equation

$$L_{k/k} = \Phi_k^T (I - W_{k+1} H_{k+1})^T L_{k+1/k+1} (I - W_{k+1} H_{k+1})\ \Phi_k \tag{62}$$

with $L_{T/T}$ defined arbitrarily depending on the problem.

Suppose we wish to investigate the effect of ΔW_k at a certain sample point of $P_{T/T}$, then we need only evaluate the matrix

$$D_{T,k}\Delta W_k\left(H_k P_{k/k-1} H_k^T + R_k\right) \Delta W_k^T D_{T,k}^T = D_{T,k}\Delta Z_{k/k} D_{T,k}^T \tag{63}$$

or the following scalar if $J = tr(L_{T/T}P_{T/T})$ is of interest:

$$tr\left[\left(\Delta W_k^T L_{k/k}\Delta W_k\right)\left(H_k P_{k/k-1}H_k^T + R_k\right)\right] = tr[L_{k/k}\Delta Z_{k/k}] \quad . \tag{64}$$

These two terms provide a valuable aid for the design of suboptimal Kalman filter gains. Suboptimal gains are of practical interest because often they may be computed simply and their use causes very little increase in the estimation error. As an example, suppose we specify an allowable ΔJ. ΔJ can be allocated among the sample points by some acceptable rules depending on the problem:

$$\Delta J = \sum_k (\Delta J)_k \quad .$$

Setting

$$\begin{aligned}(\Delta J)_k &= tr[L_{k/k}\Delta Z_{k/k}] \\ &= tr\left[\left(\Delta W_k^T L_{k/k}\Delta W_k\right)\left(H_k P_{k/k-1}H_k^T + R_k\right)\right] \quad ,\end{aligned}$$

we obtain an elliptical region around W_k, within which any ΔW_k will cause a ΔJ within the allocated $(\Delta J)_k$. In this way, a manifold may be defined around the W_k^o vs. k curve such that it contains all the allowable suboptimal gain curves. Many other ways of applying these equations are possible.

G. Applications to Problems

Depending on the particular problem, different parameters are varied and different performance criteria are used. The following are a few examples:

(1) Inaccurate statistics: Since the statistics $P_{0/-1}$, Q_{k-1}, and R_k are used to obtain the filter gain W_k, the real question is the effect of using the wrong gain. But, for computation, we may use the following equivalent problem. Let $P^*_{0/-1}$, Q^*_{k-1}, R^*_k be the erroneous statistics that give W^*_k and $P^*_{k/k}$. $P^*_{k/k}$ is not the real covariance; however, the real covariance is $P_{k/k} = P^*_{k/k} + \Delta P_{k/k}$, where $\Delta P_{k/k}$ may be computed using the variational equations [Eqs. (17) and (26)] with $\Delta W_k = 0$, $\Delta P_{0/-1} = P_{0/-1} - P^*_{0/-1}$, $\Delta Q_{k-1} = Q_{k-1} - Q^*_{k-1}$, and $\Delta R_k = R_k - R^*_k$. $P_{k/k}$ may therefore be computed from $P^*_{k/k}$, which is available in the gain computations.

(2) The value of exact statistics: Suppose we know the exact statistics, the estimation-error covariance may be improved from the $P_{k/k}$ described in (1), because we are now able to compute the exact optimum gain. Let the resultant covariance be $P^o_{k/k} = P^*_{k/k} + \Delta P^o_{k/k}$, where $\Delta P^o_{k/k}$ arises out of the $\Delta P_{0/-1}$, ΔQ_{k-1}, and ΔR_k of (1), but with $\Delta W_k = \Delta W^o_k$ instead of zero. The value of the exact statistics is therefore

$$P^o_{k/k} - P_{k/k} = P^*_{k/k} + \Delta P^o_{k/k} - P^*_{k/k} - \Delta P_{k/k} = \Delta P^o_{k/k} - \Delta P_{k/k}$$

This turns out to be the negative definite term B^o_k shown in Eq. (21).

(3) Increased sampling period: This will be discussed in detail in the next section. In general, ΔQ_{k-1}, $\Delta\Phi_{k-1}$, ΔR_k, ΔW_k exist and are parameterized by a scalar τ, the sampling period; therefore, the sensitivity matrix technique of Sec. II-C may be used. The value of ΔR_k depends in part on what, if any presmoothing of the data is used.

(4) Approximate reference model dynamics: The common practice is to represent the model error as a plant noise. The approximate dynamics are regarded as the real dynamics plus noise. The difference in performance is then between the use of the full and optimum filter and the full and optimum filter with increased plant noise: $\Delta P_{k/k}$ is computed for ΔQ_{k-1} and ΔW_k^o to give deterioration in estimator performance.

(5) A varying model parameter: Assume that the varying parameter causes a $\Delta\Phi_{k-1}$, and assume that this change is known to us so that we may incorporate it into the filter reference model as well as the filter gain. $\Delta P_{k/k}$ may be computed from Eqs. (17) and (26) with optimal B_k^o and with $\Delta\Phi_{k-1}$.

(6) Suboptimal gains: If the available nominal filter is a full optimum filter $\left(W_k^o\right)$, then $\Delta P_{k/k}$ is second order in ΔW_k through B_k. Using the error transition matrices and adjoint matrices of Sec. II-B, it is possible to specify regions around the optimal gains W_k^o such that any suboptimal gains within these regions will produce acceptable performances.

(7) Expensive observations: Suppose the estimation error at T is of major interest. We should locate the expensive observations at some point $k < T$ so that $\Delta P_{T/T}$ is as negative as possible. This information is provided by the error transition matrices and the adjoint matrices of Sec. II-B. An application of this is given in Sec. II-E.

III APPLICATIONS TO SAMPLING-PERIOD PROBLEMS

There are two types of problems concerning the sampling period. The design problem is to find a set of sampling periods such that the estimation errors are within certain specifications. The sampling periods may or may not be uniform. The real-time problem exists when, for one reason or another, the conditions are changed from the design values, and one attempts to adapt to these changes in some optimum fashion. In this section, we shall define these problems and give methodologies for their solution. An illustrative example is given in Sec. III-E.

A. Design of Uniform Sampling Period

The sampling points shown in Fig. 4 are equally spaced from 0 to T with period τ. The estimator accuracy $P_{T/T}$ at a specified time T is to meet a certain specification. An iterative procedure (Newton's method) may be defined as shown below.

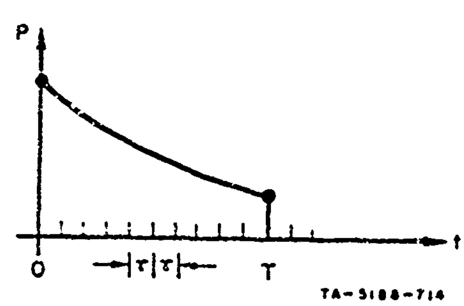

FIG. 4 UNIFORM SAMPLING PERIOD

Given:

$$P_{0/-1},\ Q,\ \Phi,\ R,\ \partial Q/\partial\tau,\ \partial\Phi/\partial\tau,\ \partial R/\partial\tau,\ L_{T/T},\ J^*$$

Find:

Sampling period τ such that

$$\mathrm{tr}(L_{T/T}P_{T/T}) \le J^* \qquad (65)$$

Algorithm:

(1) Select τ_1 as an initial sampling period, and solve the recursive equations [Eqs. (3), (46), and (47)] for $(P_{T/T})_1$ and $(\partial P_{T/T}/\partial\tau)_1$

(2) Find τ_2 by

$$(\tau_2 - \tau_1)\ \mathrm{tr}[L_{T/T}(\partial P_{T/T}/\partial\tau)_1] = J^* - \mathrm{tr}\left[L_{T/T}\left(P_{T/T}\right)_1\right] ,$$

or

$$\tau_2 = \tau_1 + \left\{J^* - \mathrm{tr}\left[L_{T/T}\left(P_{T/T}\right)_1\right]\right\}/\mathrm{tr}[L_{T/T}(\partial P_{T/T}/\partial\tau)_1] \quad . \quad (66)$$

Steps (1) and (2) are then repeated until Eq. (65) is satisfied.

The convergence of this iteration procedure remains to be proved. However, in case of difficulty one may take a sufficiently small step to at least obtain an improvement over $(P_{T/T})_1$, i.e., for $\lambda < 1$

$$\tau_2 = \tau_1 + \lambda\left\{J^* - \mathrm{tr}\left[L_{T/T}\left(P_{T/T}\right)_1\right]\right\}/\mathrm{tr}\left[L_{T/T}\left(\partial P_{T/T}/\partial\tau\right)_1\right] \quad . \quad (67)$$

B. Design of Nonuniform Sampling Period

A typical nonuniform sampling-period design problem may go like this. From results of Sec. III-A or from an examination of the computational capabilities, the total allowable number of sample points are decided. The problem is to reduce the terminal-error covariance by redistributing the sample points, while keeping the total number the same. To do this we may divide the time into m segments, and change the sampling period uniformly within each segment as depicted in Fig. 5.

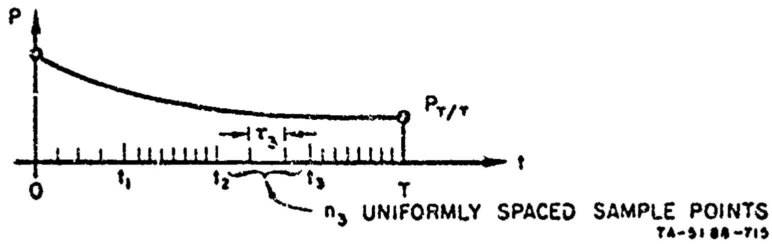

FIG. 5 NONUNIFORM SAMPLING PERIOD

Given:

$$P_{0/-1},\ Q,\ \Phi,\ R,\ \partial Q/\partial\tau,\ \partial\Phi/\partial\tau,\ \partial R/\partial\tau,\ L_{T/T} \quad . \tag{68}$$

n = number of sample points

m = number of time segments.

Find:

Sampling period, τ_i, $i = 1, 2, \ldots, m$,

$$\sum_i n_i = n,\ n_i = \frac{t_i - t_{i-1}}{\tau_i} \tag{69}$$

such that

$$J = tr(L_{T/T}P_{T/T}) \tag{70}$$

is minimized.

Discussion:

Consider τ_i, $i = 1, 2, \ldots, m$, as m scalar parameters. The following sensitivity indices may be computed using the results of Sec. II.

$$s_i = \frac{\partial tr(L_{T/T}P_{T/T})}{\partial\tau_i} \quad , \qquad i = 1, 2, \ldots, m \quad . \tag{71}$$

The problem is then equivalent to first order to the following problem:

$$\underset{\Delta\tau_i}{\text{Minimize}} \sum s_i \Delta\tau_i \quad ,$$

$$\sum_i \left(\frac{1}{\tau_i + \Delta\tau_i}\right)(t_i - t_{i-1}) = n \quad . \tag{72}$$

We must realize that the sensitivity as defined here gives only an indication of the change in performance to first order, and the results will be inaccurate if $\Delta\tau_i$ are too large.

We shall therefore limit the step sizes for $\Delta\tau_i$ by adding the following linearity constraint to Eq. (72):

$$\frac{|\Delta\tau_i|}{\tau_i} \ll 1 \quad . \tag{73}$$

From Eq. (72), it follows that a <u>necessary condition for optimum distribution points is</u> obviously

$$s_1 = s_2 = \ldots = s_m \quad ; \tag{74}$$

if not, we may always obtain an improvement in performance (a reduction in J) by moving some sample point from a segment with higher s_i. A gradient procedure is shown below for solution of this problem.

<u>Algorithm:</u>

(1) Compute the sensitivity indices s_i by Eqs. (52) and (54).

(2) Select two time segments u and ℓ that have respectively the min (s_i) and max (s_i). $n_{u\ell}$ points are then moved from the u-th segment to the ℓ-th segment, until constraint [Eq. (73)] is met for $i = u$ or $i = \ell$.

(3) If the linearity constraint is reached for segment ℓ, further improvement may be possible by moving points

from the u-th segment to the segment with the next highest s_i. In this case, repeat steps (1) and (2) until all segments are modified or until the u-th segment has reached the constraint [Eq. (73)].

(4) Recompute the s_i using the new sampling times, as determined above and proceed to step (2).

(5) Terminate the process when the s_i are sufficiently close to being identical. The examples in Sec. III-E illustrate this technique.

The sensitivity indices may of course be used in other types of gradient procedures.

C. Real-Time Allocation of Computing Resources to Multiple Trajectories

Here we address ourselves to the problem of optimum allocation of limited computation facilities to multiple trajectories when the number of trajectories cannot be predetermined accurately. We note that in case the number may indeed be predetermined, then the algorithms in Secs. III-A and III-B may be used, with slight modifications, to obtain optimum sampling schedules for the multiple trajectories. Although many types of problems may be formulated, we shall consider the following special formulation:

(1) Any one trajectory is to be filtered a fixed length of time. The terminal estimation errors are of interest.

(2) The initiation and termination times for each trajectory are arbitrary. They are therefore likely to be staggered in time to amounts that cannot be prespecified.

(3) The filtering system has available a nominal sampling rate c_i for the i-th trajectory.

(4) The computation facility capacity may be characterized by $\sum_i c_i$. That is, the facility can only handle the load

$$\sum_i c_i \leq \ell \quad , \tag{75}$$

ℓ a number denoting the capacity.

(5) The allowable real-time modifications are Δc_i.

(6) The criterion for the modification is equal degradation at terminations of each trajectory

$$\Delta j^{(1)} = \Delta j^{(2)} = \ldots \tag{76}$$

where

$$j^{(i)} = \mathrm{tr}\left[L_{T/T} P^{(i)}_{T/T}\right].$$

The solution is simple if the system has in its computer storage a proper set of sensitivity figures as follows: At real-time t, the i-th trajectory will be in a certain filtering stage denoted k_i, see Fig. 6. Let $s^{(i)}_{k_i}$ be the sensitivity of $J^{(i)}$ with respect to c_i, when c_i is changed by Δc_i from k_i to T. Then Δc_i, obtained by solving the following set of equations, is the answer to our problem:

$$\begin{gathered} s^{(1)}_{k_1} \Delta c_1 = s^{(2)}_{k_2} \Delta c_2 = \ldots \\ \sum_i (c_i + \Delta c_i) = \ell \quad . \end{gathered} \tag{77}$$

The solution is

$$\begin{gathered} \Delta c_1 = \left(\ell - \sum_i c_i\right) \Big/ \left[s^{(1)}_{k_1} \sum_i \frac{1}{s^{(i)}_{k_i}}\right] \\ \Delta c_i = s^{(1)}_{k_1} \Delta c_1 / s^{(i)}_{k_i} \quad . \end{gathered} \tag{78}$$

The $s^{(i)}_{k_i}$ are easily precomputed using the equations in Sec. II-E for sensitivity indices. The required storage

for $s_{k_i}^{(i)}$ depends a great deal on the particular problem. If various possible trajectories have similar sensitivities, then one set of s_k, $k = 0, 1, \ldots, T$ is sufficient. Further simplification may be effected by segmentation of the time axis, and assigning a sensitivity to each time segment (see Fig. 6).

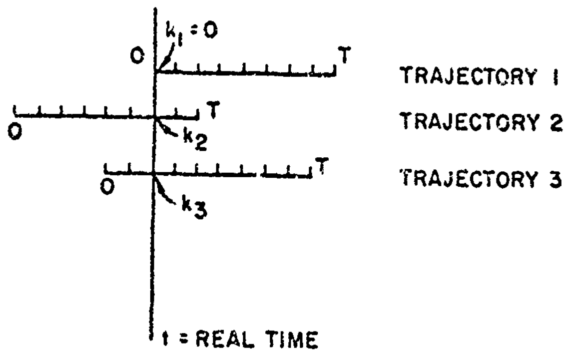

FIG. 6 RELATIONSHIP OF ITERATION STAGES TO REAL TIME

D. Real-Time Maneuver Compensation

A vehicle in motion is said to perform an unknown maneuver when its trajectory suddenly deviates a large amount from the trajectory generated by the filter reference dynamics. Unless we know when or how the vehicle will maneuver, we are forced to use in the filter a model of the dynamic equations that does not account for the maneuvers. This causes high estimation errors during unknown maneuvers.

The maneuvers can be modeled as plant noise whose magnitude and occurrence are inferred from observation; experience shows that the estimation can be improved by adjusting the filter gains closer to optimal for the increased plant noise. Our experience also shows that the speed of adaptation to maneuvers, even with the stated refinements, is limited by the sampling period. Three or four sampling periods are necessary to correct the large maneuver-induced estimation error. The direction for

further improvement of filter performance lies obviously in increased sampling rate when a maneuver is detected.

The required change in the sampling period $\Delta\tau$ may be computed from the inferred acceleration noise increase Δa by employing the sensitivity indices of Sec. II-E. We shall decrease the sampling period to counteract the effects of increased acceleration error, which leads to the following equation:

$$\Delta\tau_k = -\left(\frac{\partial J_{k+1}}{\partial \tau_k}\right)^{-1} \frac{\partial J_{k+1}}{\partial a_k} \Delta a_k \quad , \tag{79}$$

where $\partial J_{k+1}/\partial\tau_k$ and $\partial J_{k+1}/\partial a_k$ are precomputed sensitivity indicies. If J_{k+1} is defined as the root-sum-squared position error, then $\partial J_{k+1}/\partial\tau_k$ is the change in J_{k+1} due to a change in sampling time from k to k + 1; $\partial J_{k+1}/\partial a_k$ is the change due to acceleration changes from k to k + 1.

Instead of the one-stage sensitivity, we may use the multistage sensitivities $\partial J_{k+m}/\partial\tau_k$, $\partial J_{k+m}/\partial a_k$, which are the changes in root-sum-squared position error m stages ahead of the sample point k when the maneuver is detected. Thus, assuming that the maneuver acceleration uncertainty (Δa) is maintained for the m stages, we have

$$\Delta\tau = -\left(\frac{\partial J_{k+m}}{\partial \tau}\right)^{-1} \frac{\partial J_{k+m}}{\partial a} \Delta a \quad , \tag{80}$$

where the sensitivity indices are obtained by solving the recurrent equations [Eqs. (46), (47), (50), and (34)] m times using zero initial values.

E. Examples

The procedures discussed in Secs. III-A and -B will now be illustrated by a simple estimator problem using a $1/s^2$ plant. The numerical values used are, however, on the order of what one might encounter in the estimation of ballistic trajectories.

Example 1:

Assume that the ballistic vehicle moves in a straight line with no other driving force than the acceleration noises. That is, let p be the

position of the vehicle and n_a the acceleration noise, then

$$\ddot{p} = n_a \quad . \tag{81}$$

The position is observed in the presence of additive noise n_m:

$$z = p + n_m \quad . \tag{82}$$

The discrete-time, state-variable description is

$$\begin{aligned} x_{k+1} &= \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} x_k + w_k \\ z_k &= [1 \quad 0]\, x_k + v_k \quad , \end{aligned} \tag{83}$$

where T is the sampling interval from k to k + 1 and

$$x_k = \begin{bmatrix} p_k \\ \dot{p}_k \end{bmatrix} \quad . \tag{84}$$

The quantities w_k and v_k are the plant noise and measurement noise with covariance matrices Q_k and R_k, respectively, and x_0 is assumed to be Gaussian-distributed with mean $E(x_0)$ and covariances $P_{0/-1}$.

The Kalman filter, Eq. (2), computes $\hat{x}_{k/k}$, the estimate of x_k, using the observations z_0, z_1, ..., z_k.

We shall assume the following numerical values:

$$R_k = 10^4 \quad , \tag{85}$$

for 100 ft standard deviation in measurement noise,

$$Q_k = \begin{bmatrix} 0.259\tau & 0 \\ 0 & 104\tau \end{bmatrix}$$

for 1-g acceleration error,

$$Q_k = \begin{bmatrix} 25.9\tau & 0 \\ 0 & 10400\tau \end{bmatrix} \qquad (86)$$

for 10-g acceleration error, and

$$P_{0/-1} = \begin{bmatrix} 10^6 & 0 \\ 0 & 10^6 \end{bmatrix}$$

for initial position and velocity standard deviations of 1000 ft and 1000 ft/sec.

τ varies in the program, with 0.1 sec as its nominal value.

The total time of the estimation process is 5 sec, which is divided into five segments of 1 sec each. Within each segment, the sampling points are equally spaced. Thus, when the number of sampling points are changed from 10 to 8 in a certain segment, the sampling period is changed from 1/10 sec to 1/8 sec.

Case 1: Plant Noise is 1 g

First we use the (10, 10, 10, 10, 10) sampling scheme--numbers in the parentheses denote the number of sampling points in each segment. The position estimation-error covariance is shown in curve a of Fig. 7, where a terminal position error covariance of 876 ft^2 is observed.

Now, suppose the computation capability of the system is such that only 40 sample points are allowable in the 5-sec interval instead of the proposed 50. First, we modify the filter in the most obvious way, namely, using the scheme (8, 8, 8, 8, 8). The results are shown in curve b of Fig. 7; the terminal error has deteriorated 21 percent to 1060 ft^2. We will now use a better sampling scheme in order to decrease the terminal error. For this, we need the guidance of a sensitivity computation.

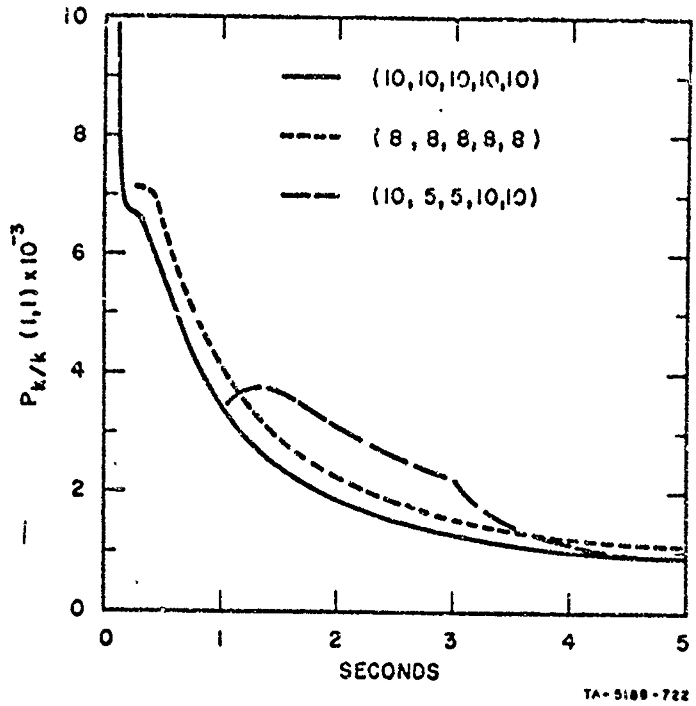

FIG. 7 POSITION COVARIANCES, 1-g PLANT NOISE

The sensitivity of terminal error to sampling period, $s(\tau_i)$, is plotted in Fig. 8. The least sensitivity occurs in the second segment. The reduction in sampling points, therefore, should be in the second

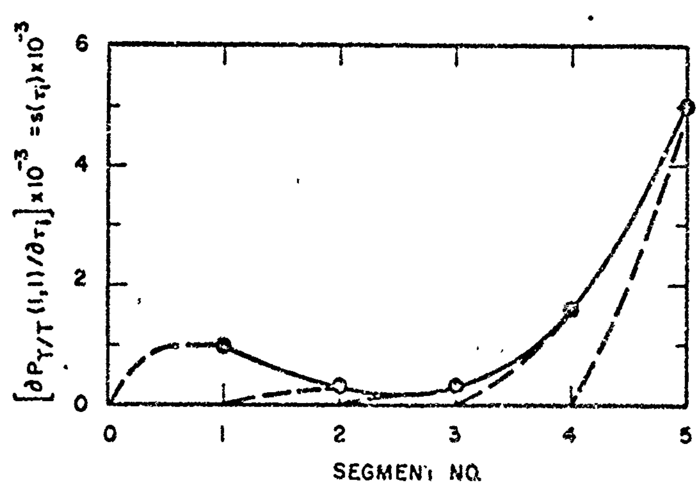

FIG. 8 SENSITIVITY OF TERMINAL ERROR TO SAMPLING PERIOD, 1-g PLANT NOISE

segment. Allowing a reduction of 50 percent for linearity as discussed in Sec. III-B, the number of sampling points is reduced within the second segment from 10 to 5. Since the next lower sensitivity is in segment 3, the remaining reduction is carried out in segment 3; from 10 to 5. With the new scheme of (10, 5, 5, 10, 10), the terminal position error covariance change, from the sensitivity indices, should be:

$$\Delta P = s(\tau_2) \times \Delta\tau_2 + s(\tau_3) \times \Delta\tau_3$$

$$= 227 \times (0.2 - 0.1) + 331 \times (0.2 - 0.1) = 55.8 \text{ ft}^2 \quad .$$

The error covariance with (10, 5, 5, 10, 10) is computed and plotted in curve c of Fig. 7. The terminal position-error covariance of 894 ft^2 represents an increase of only 18 ft^2 (compare with 55.8 ft^2 as predicted above) over the scheme using (10, 10, 10, 10, 10). It is interesting to note that 10 sample points may be removed from the right places with very little effect on the terminal estimation accuracy. The only additional labor involved is the computation of the sensitivity indices $s(\tau_i)$.

The diagonal elements of the adjoint matrices are plotted in Fig. 9. The curve a shows the (1, 1) element of $L_{k/k}$ with $L_{T/T} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, which may be interpreted as the sensitivity of the terminal position-error covariance to the k-th position-error covariance. Similarly, curve b shows the sensitivity with respect to the k-th velocity error covariance; i.e., the (2, 2) element of $L_{k/k}$. This plot provides additional insight into the estimation problem beyond those given by the sensitivity indices. The design of an estimator may be greatly improved with this insight. Referring again to Fig. 9, it can be seen that in the 4-th segment (3 to 4 sec), velocity-error sensitivity predominates. This indicates that if we perform a velocity measurement at that time, terminal estimation accuracy may be much improved. To obtain some definite quantities, the velocity measurement is assumed to have an error covariance of 100 $(ft/sec)^2$. We shall perform the velocity measurement at 3.5 sec, at which time the nominal velocity covariance is 413 (ft/sec)

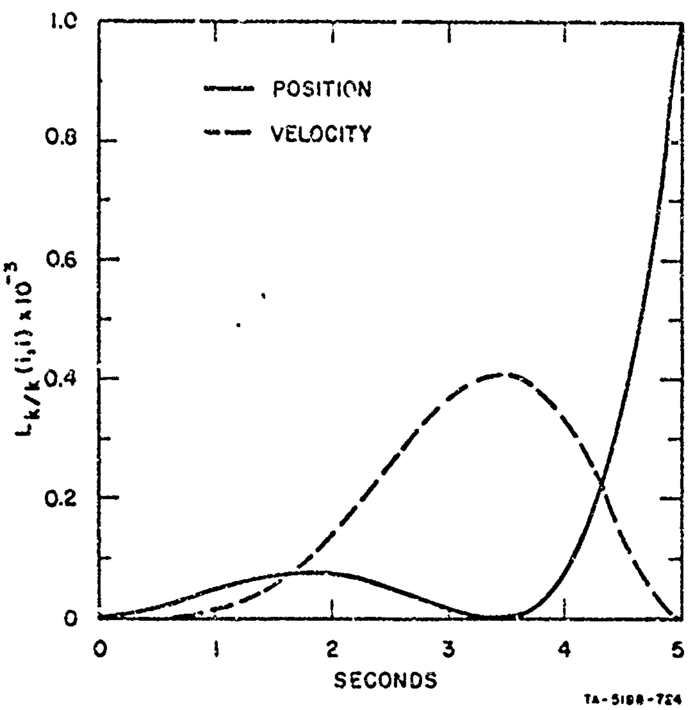

FIG. 9 TERMINAL POSITION SENSITIVITY TO COVARIANCE AT TIME k, 1-g PLANT NOISE

from computations. With a single measurement, the velocity estimate improvement is at least 413 - 100 = 313 $(ft/sec)^2$. Since curve b in Fig. 9 shows a sensitivity of 0.4 at 3.5 sec, the terminal position accuracy is expected to improve by 0.4 × 313 = 125 ft^2--an improvement of 125/876 = 14 percent.

Further insight that may be obtained from the adjoint matrix of Fig. 6 is that although position measurements are made in the 3 to 4-sec segment, their predominant purpose is to obtain better velocity estimates, which in turn improves the terminal accuracy. Given a choice of position or velocity measurement improvement in 3 to 4 sec, one should take the velocity improvement if Fig. 9 is applicable to the system.

<u>Case 2: Plant Noise is 10 g</u>

With the plant noise increased, we would expect that what the estimator does initially has less effect on the terminal estimation error. This fact is dramatically illustrated in the following.

The position-error covariance is shown in curve a of Fig. 10 for a (10, 10, 10, 10, 10) sampling scheme. The terminal position-error covariance is 2249 ft^2. The terminal-error covariance sensitivities to sampling period are shown in Fig. 11. It is seen that changing the

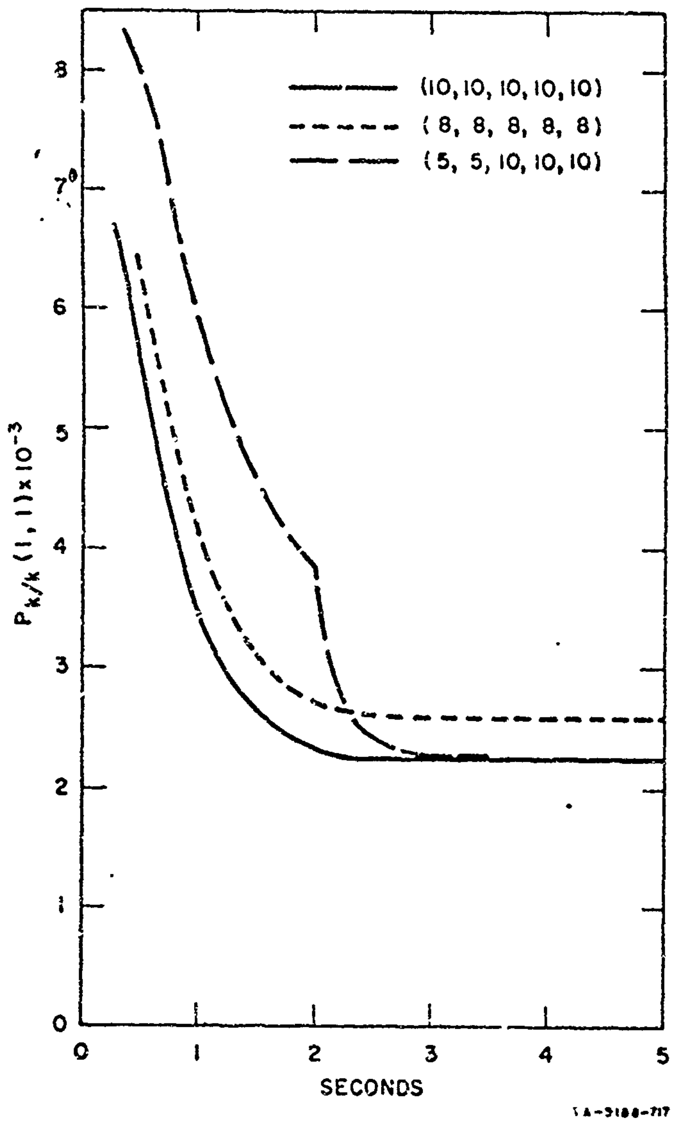

FIG. 10 POSITION COVARIANCES, 10-g PLANT NOISE

sampling period in the first two segments will have virtually no effect on the terminal error. Therefore, the reduction in sampling period should be confined to the first two segments, 0 to 2 sec. According to

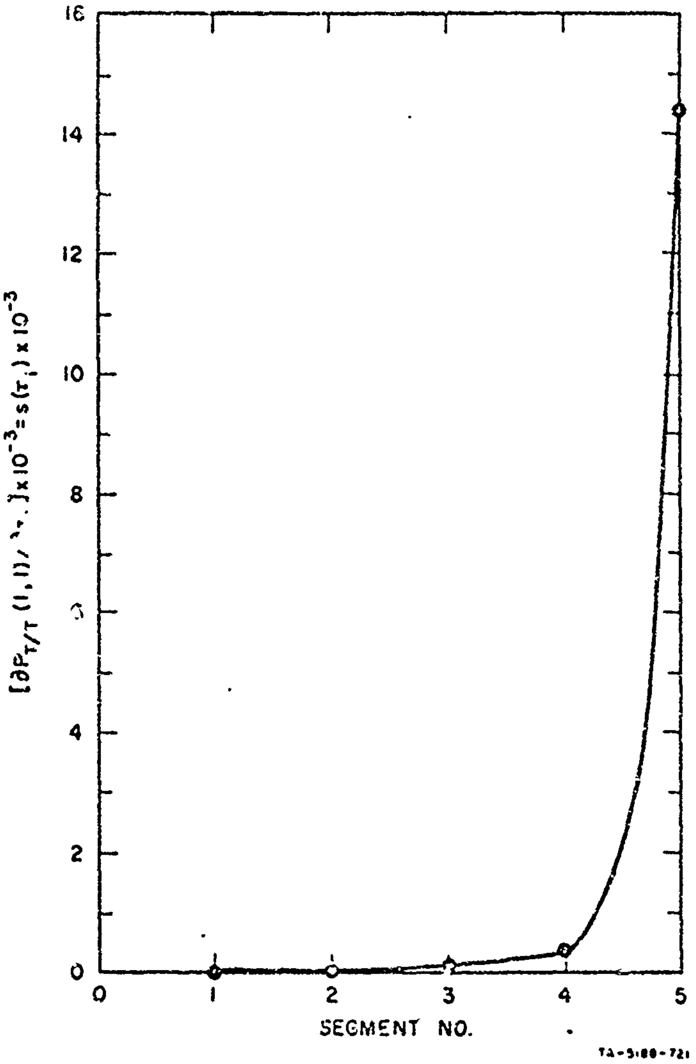

FIG. 11 SENSITIVITY OF TERMINAL ERROR TO SAMPLING PERIOD, 10-g PLANT NOISE

this argument, a (5, 5, 10, 10, 10) scheme is tried with the results plotted in curve c of Fig. 10. The terminal error of 2250 ft^2, as compared with 2249 ft^2, is almost exactly the same as the (10, 10, 10, 10, 10) scheme. If the same 10-point reduction were taken uniformly. namely the (8, 8, 8, 8, 8) scheme, the terminal error would be 2600 ft^2 as shown in curve b of Fig. 10--a 15.5-percent increase.

Figure 12 shows the sensitivity of the terminal position covariance to the k-th covariance (diagonal elements of the adjoint matrix $L_{k/k}$).

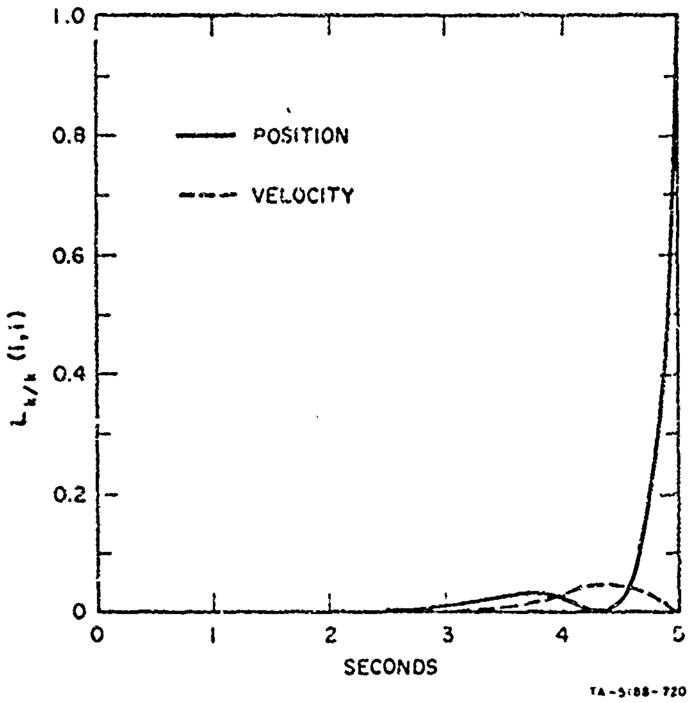

FIG. 12 TERMINAL POSITION SENSITIVITY TO COVARIANCE, 10-g PLANT NOISE

Similar to the 1-g case, we see from this plot the advantage of making a velocity measurement at 4.3 sec. Assuming a measurement covariance of 100 $(ft/sec)^2$, the terminal position covariance is expected to decrease by approximately the amount

$$L_{k/k}(2,2) \times [P_{k/k}(2,2) - 100] = (0.05) \times (8240 - 100) = 407 \; (ft^2) \quad ,$$

which represents a 407/2249 = 18 percent decrease.

Example 2:

In this example, we shall point out the precautions that should be exercised in deriving sensitivity indices for smoothed measurements. Under certain conditions the sampling-period sensitivity may become negative, indicating one should update less frequently in order to reduce terminal error. While this may seem to contradict our engineering intuition, mathematically it is perfectly rigorous. Nevertheless, if

this becomes unacceptable for any reason, it is essential to analyze carefully the assumptions concerning the measurement smoothing and the plant noise.

In Example 1, assume the measurements are smoothed such that the measurement covariance is

$$R_k = \frac{10^3}{\tau} \text{ ft}^2 \quad . \tag{87}$$

Note that when $\tau = 0.1$ sec, the nominal sampling period, R_k is 10^4 ft^2, the value used in Example 1. In Fig. 13, the sensitivities $s(\tau_i)$ are plotted, for plant noises of 1 g and 10 g. Both cases show negative sensitivities toward the end of the estimation process, meaning that the terminal accuracy is improved for more widely separated sample points.

To verify the negative sensitivity, the covariance equation is recomputed for the 10-g case with the number of sample points reduced from

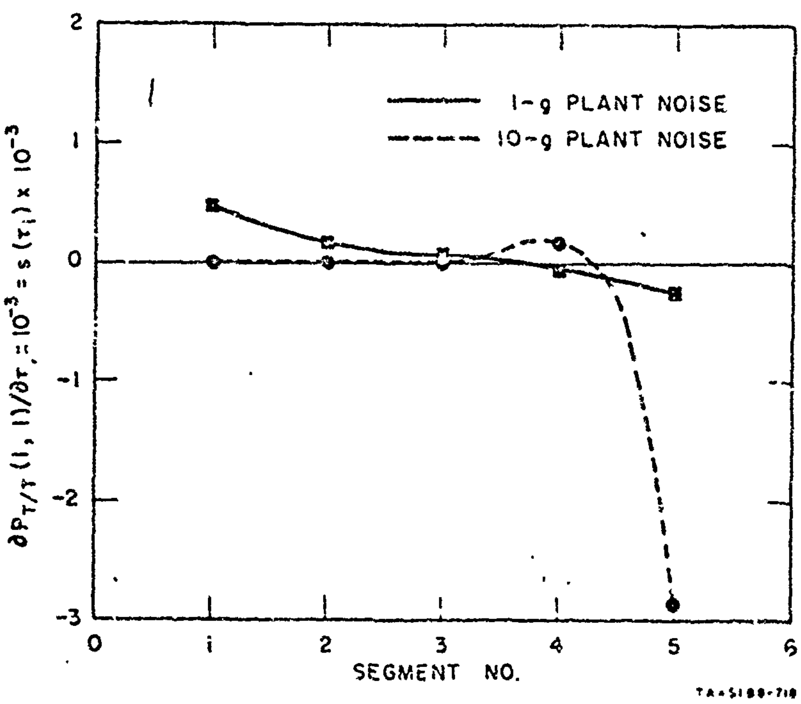

FIG. 13 SENSITIVITY OF TERMINAL ERROR TO SAMPLING PERIOD, SMOOTHED MEASUREMENTS

10 to 6 in the last segment. The new terminal position covariance is 2070 ft^2, a reduction of 179 ft^2 from the original 2249 ft^2.

The negative sensitivity arises from the fact that the reduction in measurement covariance, Eq. (87), outraces the increase in plant noise, Eq. (86), as sampling period increases. We shall now examine the validity of these assumptions about the noise.

It may be shown that if the plant noise is approximated by a white noise in the continuous-time description of the dynamical system, then its discrete description will have a noise covariance, as described by Eq. (86), proportional to τ. However, if the plant noise is meant to represent mathematical modeling errors, this may not be a valid assumption. A better noise model is perhaps an acceleration noise that remains constant over τ, where its magnitude is a random variable, zero mean and Gaussian in distribution with standard deviation a. For a given time interval τ, the plant noise covariance matrix would then be better expressed as

$$Q = \begin{bmatrix} \left(1/2\ a\tau^2\right)^2 & 0 \\ 0 & (a\tau)^2 \end{bmatrix} \quad . \tag{88}$$

The fourth and second powers of τ here indicate that as the sampling period τ increases, the true plant noise is likely to increase more rapidly than those assumed in Eq. (86). This will cause the sampling period sensitivities to tend towards positive values.

Returning to the question of the measurement noise covariance, we see that the form of Eq. (87) is valid only if the data are "perfectly" smoothed. Namely, the data must be fitted to a curve representing the precise output of the dynamical system. Yet in practice, data are smoothed using arbitrary curves such as straight lines or polynomials. In doing so, additional errors are introduced beyond those represented by Eq. (87). A more truthful representation of the smoothed measurement is perhaps

$$R_k = \frac{10^3}{\tau^\alpha} \quad , \qquad \alpha < 1 \quad .$$

The effect is that as sampling period τ increases, the noise covariance does not decrease as rapidly. This will also tend to move the sensitivity to the positive side.

We have seen that any refinement of our noise assumptions tends to move away from a negative sensitivity index for sampling period. It is therefore doubtful that the negative sensitivity should be taken seriously as a significant phenomenon.

IV ONE-STAGE TRADE-OFF STUDIES USING A SIMPLIFIED MODEL

A. One-Stage Trade-Off

The purpose of this section is to provide formulas and curves that will promote a general understanding of the relationship between filter performance and various parameters, and the resulting trade-offs between these parameters. Such understanding is necessary when one is doing, for example, trade-off studies between sampling period and plant noise, sampling period and measurement accuracy, plant noise and measurement accuracy, etc.

Since our interest is in deriving qualitative relationships that are easily computable, we shall limit our discussion to the following case. These assumptions should be kept in mind when the results are applied to practical problems.

(1) $1/s^2$ plant: The state equation is

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + u \quad , \quad \text{or} \quad x_{k+1} = \begin{bmatrix} 1 & \tau^* \\ 0 & 1 \end{bmatrix} x_k + u_k \quad , \tag{89}$$

where x is a 2-vector composed of position and velocity, u is the plant noise, τ^* is the sampling period.

(2) Position measurement: The observation equation is

$$z_k = Hx_k + v_k = [1 \quad 0]\, x_k + v_k \quad , \tag{90}$$

where v is the observation noise, $E[v_k] = 0$, $E[v_k^2] = R_k$.

(3) Plant noise:

$$E[u_k] = 0$$

$$E[u_k\ u_k^T] = Q_k = \begin{bmatrix} q_x & 0 \\ 0 & q_v \end{bmatrix} \quad . \tag{91}$$

(4) u_k, v_k, $k = 0, 1, 2, \ldots$, are Gaussian white and uncorrelated.

The formula we shall develop is for the covariance of the estimation errors from k to k + 1 sampling point. They apply, therefore, only to local considerations. The behavior of the error covariance matrix is depicted symbolically in Fig. 14.

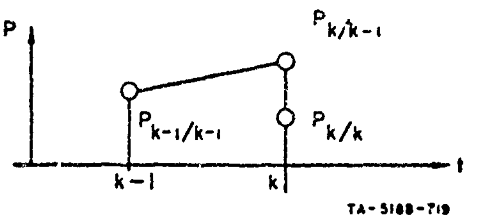

FIG. 14 PREDICTION AND UPDATING STAGES IN KALMAN FILTER

From $P_{k-1/k-1}$ to $P_{k/k-1}$, prediction of the state is performed; the covariance usually grows because of the velocity error and the plant noise. From $P_{k/k-1}$ to $P_{k/k}$, the k-th measurement is used to update the state; the reduction in the covariance is a function of the measurement accuracy.

Let

$$P_{k-1/k-1} = \begin{bmatrix} 1 & \left(p_{xv}\right)_{k-1/k-1} \\ \left(p_{xv}\right)_{k-1/k-1} & 1 \end{bmatrix} \tag{92}$$

$$P_{k/k-1} = \begin{bmatrix} \left(p_x\right)_{k/k-1} & \left(p_{xv}\right)_{k/k-1} \\ \left(p_{xv}\right)_{k/k-1} & \left(p_v\right)_{k/k-1} \end{bmatrix} \tag{93}$$

$$P_{k/k} = \begin{bmatrix} \left(p_x\right)_{k/k} & \left(p_{xv}\right)_{k/k} \\ \left(p_{xv}\right)_{k/k} & \left(p_v\right)_{k/k} \end{bmatrix} \quad . \tag{94}$$

Note that the diagonal elements of $P_{k-1/k-1}$ in Eq. (92) are unity. $P_{k-1/k-1}$ may always be brought to this form by a normalization process which is discussed in Appendix C.

The prediction process, $P_{k-1/k-1}$ to $P_{k/k-1}$, is given by the equation

$$P_{k/k-1} = \Phi_{k-1} \, P_{k-1/k-1} \, \Phi^T_{k-1} + Q_{k-1} \quad . \tag{95}$$

Under our assumptions, we have

$$\begin{aligned} \left(p_x\right)_{k/k-1} &= 1 + 2 \left(p_{xv}\right)_{k-1/k-1} \tau + \tau^2 + q_x \\ \left(p_{xv}\right)_{k/k-1} &= \left(p_{xv}\right)_{k-1/k-1} + \tau \quad . \\ \left(p_v\right)_{k/k-1} &= 1 + q_v \quad . \end{aligned} \tag{96}$$

The updating process, $P_{k/k-1}$ to $P_{k/k}$, is given by the following equation, assuming the optimal gain is used:

$$P_{k/k} = P_{k/k-1} - P_{k/k-1} H^T \left(HP_{k/k-1} H^T + R\right)^{-1} HP_{k/k-1} \quad . \tag{97}$$

Using the assumptions, we have

$$\left(p_x\right)_{k/k} = \frac{\left(p_x\right)_{k/k-1}}{\dfrac{\left(p_x\right)_{k/k-1}}{R_x} + 1}$$

$$\left(p_{xv}\right)_{k/k} = \frac{\left(p_{xv}\right)_{k/k-1}}{\dfrac{\left(p_x\right)_{k/k-1}}{R_x} + 1} \tag{98}$$

$$\left(p_v\right)_{k/k} = \frac{\left(p_v\right)_{k/k-1} + \dfrac{\left(p_x\right)_{k/k-1}\left(p_v\right)_{k/k-1} - \left(p_{xv}\right)^2_{k/k-1}}{R_x}}{\dfrac{\left(p_x\right)_{k/k-1}}{R_x} + 1}$$

$$= \left(p_v\right)_{k/k-1} - \frac{\left(p_{xv}\right)^2_{k/k-1}}{R_x + \left(p_x\right)_{k/k-1}}$$

$P_{k/k}$ may be expressed directly in terms of $P_{k-1/k-1}$ by combining Eqs. (96) and (98):

$$\left(p_x\right)_{k/k} = \frac{1 + 2\left(p_{xv}\right)_{k-1/k-1}\tau + \tau^2 + q_x}{1 + \dfrac{\left[1 + 2\left(p_{xv}\right)_{k-1/k-1}\tau + \tau^2 + q_x\right]}{R_x}}$$

$$\left(p_{xv}\right)_{k/k} = \frac{\left(p_{xv}\right)_{k-1/k-1} + \tau}{1 + \dfrac{\left[1 + 2\left(p_{xv}\right)_{k-1/k-1}\tau + \tau^2 + q_x\right]}{R_x}} \tag{99}$$

$$\left(p_v\right)_{k/k} = 1 + q_v - \frac{\left[\left(p_{xv}\right)_{k-1/k-1} + \tau\right]^2/R_x}{1 + \dfrac{\left[1 + 2\left(p_{xv}\right)_{k-1/k-1}\tau + \tau^2 + q_x\right]}{R_x}}$$

For trade-off studies, we need the sensitivity of $P_{k/k}$ with respect to various parameters. Those are easily derived using the result (see Sec. II-C) that if $P_{k/k-1}$ and R_k are functions of a scalar parameter π and if W is chosen optimally or fixed, then

$$\left\{\begin{aligned} \frac{\partial}{\partial\pi} P_{k/k-1} &= \frac{\partial}{\partial\pi}\left(\Phi_{k-1}\, P_{k-1/k-1}\, \Phi^T_{k-1}\right) + \frac{\partial}{\partial\pi} Q_{k-1} \\ \frac{\partial}{\partial\pi} P_{k/k} &= (I - W_k H)\left(\frac{\partial}{\partial\pi} P_{k/k-1}\right)(I - W_k H)^T + W_k\left(\frac{\partial}{\partial\pi} R_k\right) W_k^T \end{aligned}\right. \tag{100}$$

$$W_k = P_{k/k-1}\, H^T \left(H\, P_{k/k-1}\, H^T + R\right)^{-1} = \text{optimal gain}$$

or

$$W_k = \text{fixed} \quad .$$

Let us choose the parameters τ, q_x, q_v, and R_x for a trade-off study. Furthermore, q_x, q_v, and R_x, and Φ are dependent on sampling period (τ), q_x and q_v on acceleration error (α), and R_x on parameter (β):

$$\begin{aligned} q_x &= q_x(\alpha, \tau) \quad , \qquad q_v = q_v(\alpha, \tau) \\ R_x &= R_x(\beta, \tau) \\ \Phi &= \Phi(\tau) \quad . \end{aligned} \tag{101}$$

The parameters are then α, β, and τ:

$$\begin{aligned} P_{k/k-1} &= P_{k/k-1}(\alpha, \tau) \\ P_{k/k} &= P_{k/k}(\beta, \tau, P_{k/k-1}) \quad . \end{aligned} \tag{102}$$

Using Eq. (96), the following equations may be written

$$\frac{\partial P_{k/k-1}}{\partial \alpha} = \begin{bmatrix} \frac{\partial q_x}{\partial \alpha} & 0 \\ 0 & \frac{\partial q_v}{\partial \alpha} \end{bmatrix}$$

$$\frac{\partial P_{k/k-1}}{\partial \tau} = \begin{bmatrix} 2\left[\tau + \left(p_{xv}\right)_{k-1/k-1}\right] + \frac{\partial q_x}{\partial \tau} & 1 \\ 1 & \frac{\partial q_v}{\partial \tau} \end{bmatrix} \tag{103}$$

$$\frac{\partial P_{k/k-1}}{\partial \beta} = 0$$

$$\frac{\partial R_x}{\partial \alpha} = 0 \quad .$$

Substituting Eq. (103) into Eq. (100), we obtain the following expressions for the elements of $\partial P_{k/k}/\partial \alpha$, the sensitivity to the plant noise parameter α:

$$\frac{\partial \left(p_x\right)_{k/k}}{\partial \alpha} = (1 - k_1)^2 \frac{\partial q_x}{\partial \alpha}$$

$$\frac{\partial \left(p_{xv}\right)_{k/k}}{\partial \alpha} = (1 - k_1)\ (- k_2)\ \frac{\partial q_x}{\partial \alpha} \tag{104}$$

$$\frac{\partial \left(p_v\right)_{k/k}}{\partial \alpha} = k_2^2 \frac{\partial q_x}{\partial \alpha} + \frac{\partial q_v}{\partial \alpha} \quad ,$$

where

$$k_1 = \frac{\left(p_x\right)_{k/k-1}}{R_x + \left(p_x\right)_{k/k-1}} \quad \text{and} \quad k_2 = \frac{\left(p_{xv}\right)_{k/k-1}}{R_x + \left(p_x\right)_{k/k-1}} \tag{105}$$

$$1 - k_1 = \frac{R_x}{R_x + \left(p_x\right)_{k/k-1}} .$$

Similarly, for the sampling period parameter τ:

$$\frac{\partial \left(p_x\right)_{k/k}}{\partial \tau} = (1 - k_1)^2 \left[2\tau + 2\left(p_{xv}\right)_{k-1/k-1} + \frac{\partial q_x}{\partial \tau}\right] + k_1^2 \frac{\partial R_x}{\partial \tau}$$

$$\frac{\partial \left(p_{xv}\right)_{k/k}}{\partial \tau} = (1 - k_1) \left\{- k_2 \left[2\tau + 2\left(p_{xv}\right)_{k-1/k-1} + \frac{\partial q_x}{\partial \tau}\right] + 1\right\} - k_1 k_2 \frac{\partial R_x}{\partial \tau} \tag{106}$$

$$\frac{\partial \left(p_v\right)_{k/k}}{\partial \tau} = k_2^2 \left[2\tau + 2\left(p_{xv}\right)_{k-1/k-1} + \frac{\partial q_x}{\partial \tau}\right] - 2k_2 + \frac{\partial q_v}{\partial \tau} + k_2^2 \frac{\partial R_x}{\partial \tau} .$$

And similarly, for the measurement noise parameter β:

$$\frac{\partial \left(p_x\right)_{k/k}}{\partial \beta} = k_1^2 \frac{\partial R_x}{\partial \beta}$$

$$\frac{\partial \left(p_{xv}\right)_{k/k}}{\partial \beta} = k_1 k_2 \frac{\partial R_x}{\partial \beta} \tag{107}$$

$$\frac{\partial \left(p_v\right)_{k/k}}{\partial \beta} = k_2^2 \frac{\partial R_x}{\partial \beta} .$$

Example 1:

Let us illustrate the use of the above equations by a problem frequently encountered in Kalman filtering: the trade-off between model accuracy and sampling period ($1/s^2$ plant approximation). In an attempt to reduce computation time, one wishes to use a simplified dynamics model. However, in doing so, the state estimate error increases. This may be compensated by sampling (processing data) more frequently, which increases the computation time. One, therefore, needs a set of alternatives on sampling-period and model accuracy to analyze their computational requirements.

Trade-off between plant noise and sampling period may also occur in the estimation of maneuvering reentry vehicle trajectories, as discussed in Sec. III-D. A maneuver may be modeled as an increase in the plant noise that causes larger than expected estimation errors if the sampling period is held fixed. By sampling more frequently, this error may converge more rapidly.

Let plant noise be represented by

$$q_x = \left(\frac{1}{2}\,\alpha\,\tau^2\right)^2 = \frac{1}{4}\,\alpha^2\,\tau^4 \quad ,$$

where τ and α are respectively the normalized sampling period and normalized acceleration standard deviation.

$$q_v = (\alpha\,\tau)^2 = \alpha^2\,\tau^2 \quad . \tag{108}$$

And let the measurement noise be independent of sampling period. Then,

$$\frac{\partial R_x}{\partial \tau} = 0$$

$$\frac{\partial q_x}{\partial \alpha} = \frac{1}{2}\,\alpha\,\tau^4 \quad , \quad \frac{\partial q_x}{\partial \tau} = \alpha^2\,\tau^3 \tag{109}$$

$$\frac{\partial q_v}{\partial \alpha} = 2\,\alpha\,\tau^2 \quad , \quad \frac{\partial q_v}{\partial \tau} = 2\alpha^2\,\tau \quad .$$

Let $\Delta\alpha$ represent the increase in plant noise from normal to high; $\Delta\tau$, the required change in sampling period to bring the position estimate error back to its original value, is obtained from the first equations of Eqs. (104) and (105):

$$\frac{\partial\left(p_x\right)_{k/k}}{\partial\alpha}\,\Delta\alpha = -\frac{\partial\left(p_x\right)_{k/k}}{\partial\tau}\,\Delta\tau$$

(110)

$$(1-k_1)^2\,\frac{\partial q_x}{\partial\alpha}\,\Delta\alpha = \left\{(1-k_1)^2\left[2\tau + 2\left(p_{xv}\right)_{k-1/k-1} + \frac{\partial q_x}{\partial\tau}\right] + k_1^2\,\frac{\partial R_x}{\partial\tau}\right\}(-\Delta\tau)$$

$$\frac{1}{2}\,\alpha\,\tau^4\,(\Delta\alpha) = \left[2\tau + 2\left(p_{xv}\right)_{k-1/k-1} + \alpha^2\,\tau^3\right](-\Delta\tau)$$

$$(\Delta\tau) = -\frac{\frac{1}{2}\,\alpha\,\tau^4}{2\tau + 2\left(p_{xv}\right)_{k-1/k-1} + \alpha^2\,\tau^3}\,(\Delta\alpha) \qquad (111)$$

$$\frac{\Delta\tau}{\tau} = -\frac{1}{2}\,\frac{1}{1 + \frac{2}{\alpha^2\,\tau^3}\left[\tau + \left(p_{xv}\right)_{k-1/k-1}\right]}\left(\frac{\Delta\alpha}{\alpha}\right) \quad . \qquad (112)$$

When $\alpha^2\,\tau^3$ is small, $\Delta\tau/\tau \cong 0$; when large, $\Delta\tau/\tau \cong -1/2\ \Delta\alpha/\alpha$. In the maneuvering reentry vehicle example above, this equation may be used to change the sampling period by the amount $\Delta\tau$ after estimating the noise increase $\Delta\alpha$. $\Delta\alpha$ may be inferred from the measurement residuals.

It should be noted that Eq. (109) is given in normalized quantities; see Appendix C for reconversion into the original variables.

Example 2:

The trade-off between measurement noise and sampling period ($1/s^2$ plant approximation) will now be considered. In radar observation of reentry vehicles, this question may come up in different phases of design and operation. In the design phase, there is the question of proper balance between radar accuracy and data processing requirement. For example, if the radar noise is reduced, the sampling period may be increased--meaning less data processing. The measurement noise may also change during operation; for example, damage to the radar elements, or splitting of the array into subarrays for simultaneous observation of multiple targets will, in general, increase the noise. This may be compensated by faster sampling.

To compute the change in sampling period necessary to compensate for the changed measurement noise, we may use the following formula:

$$\frac{\partial\left(p_x\right)_{k/k}}{\partial\beta}\,\Delta\beta = -\frac{\partial\left(p_x\right)_{k/k}}{\partial\tau}\,\Delta\tau \tag{113}$$

which becomes, upon using Eqs. (106) and (107), for a local correction

$$k_1^2\,\frac{\partial R_x}{\partial\beta}\,\Delta\beta = -\left\{(1-k_1)^2\left[2\tau + 2\left(p_{xv}\right)_{k-1/k-1} + \frac{\partial q_x}{\partial\tau}\right] + k_1^2\,\frac{\partial R_x}{\partial\tau}\right\}\Delta\tau \quad . \tag{114}$$

If the plant noise is $q_x = 1/4\ \alpha^2\ \tau^4$ and the measurement noise is

$$R_x = \beta/\tau \quad , \tag{115}$$

which represents respectively model acceleration error and presmoothed measurements, Eq. (114) may be rewritten as

$$\frac{(\Delta\beta)}{\tau} = -\left\{\left[\frac{R_x}{\left(p_x\right)_{k/k-1}}\right]^{-2}\left[2\tau + 2\left(p_{xv}\right)_{k-1/k-1} + \alpha^2\ \tau^3\right] - \frac{\beta}{\tau^2}\right\}\Delta\tau \tag{116}$$

$$\frac{\Delta\tau}{\tau} = -\frac{1}{-1 + \dfrac{(\tau R_x)^2}{\beta}\,\dfrac{2\tau + 2\left(p_{xv}\right)_{k-1/k-1} + \alpha^2\ \tau^3}{1 + 2\left(p_{xv}\right)_{k-1/k-1}\tau + \tau^2 + q_x}}\,\frac{\Delta\beta}{\beta} \quad , \tag{117}$$

when τ^2/β is large, $\Delta\tau/\tau \cong 0$; when small, $\Delta\tau/\tau \cong \Delta\beta/\beta$.

Normalized variables are used in Eq. (117); Appendix C gives equations for reconversion to the real variables.

B. One-Stage Sampling-Period Sensitivities

In this section, plots are presented giving the sensitivity and the reduction of estimation-error covariance, in one stage of Kalman filtering, as functions of sampling period for various types of plant and measurement noises. Using these plots, the effect of sampling-period variations to local-error variations may be estimated quickly. Their

effects on errors at later times can be estimated using the adjoints, Sec. II-B. The dynamical system is a double integrator [see Eq. (99)] subject to random acceleration excitation; the measured quantity is position [see Eq. (90)]. Normalized quantities are used. Appendix C shows the relationships between the actual values (denoted by a super *) and the normalized quantities (unstarred quantities) to be as follows:

$$P^*_{k-1/k-1} = \begin{bmatrix} \left(p^*_x\right)_{k-1/k-1} & \left(p^*_{xv}\right)_{k-1/k-1} \\ \left(p^*_{xv}\right)_{k-1/k-1} & \left(p^*_v\right)_{k-1/k-1} \end{bmatrix} \tag{118}$$

and

$$P_{k-1/k-1} = \begin{bmatrix} \left(p_x\right)_{k-1/k-1} & \left(p_{xv}\right)_{k-1/k-1} \\ \left(p_{xv}\right)_{k-1/k-1} & \left(p_v\right)_{k-1/k-1} \end{bmatrix}$$
$$= \begin{bmatrix} 1 & \left(p^*_{xv}\right)_{k-1/k-1} \Big/ \sqrt{\left(p^*_x\, p^*_v\right)_{k-1/k-1}} \\ \left(p^*_{xv}\right)_{k-1/k-1} \Big/ \sqrt{\left(p^*_x\, p^*_v\right)_{k-1/k-1}} & 1 \end{bmatrix} \tag{119}$$

$$P_{k/k-1} = \begin{bmatrix} \left(p_x\right)_{k/k-1} & \left(p_{xv}\right)_{k/k-1} \\ \\ \left(p_{xv}\right)_{k/k-1} & \left(p_v\right)_{k/k-1} \end{bmatrix}$$

$$= \begin{bmatrix} \left(p_x^*\right)_{k/k-1} \Big/ \left(p_x^*\right)_{k-1/k-1} & \left(p_{xv}^*\right)_{k/k-1} \Big/ \sqrt{\left(p_x^* \; p_v^*\right)_{k-1/k-1}} \\ \\ \left(p_{xv}^*\right)_{k/k-1} \Big/ \sqrt{\left(p_x^* \; p_v^*\right)_{k-1/k-1}} & \left(p_v^*\right)_{k/k-1} \Big/ \left(p_v^*\right)_{k-1/k-1} \end{bmatrix} \tag{120}$$

$$P_{k/k} = \begin{bmatrix} \left(p_x\right)_{k/k} & \left(p_{xv}\right)_{k/k} \\ \\ \left(p_{xv}\right)_{k/k} \Big/ & \left(p_v\right)_{k/k} \end{bmatrix}$$

$$= \begin{bmatrix} \left(p_x\right)_{k/k} \Big/ \left(p_x^*\right)_{k-1/k-1} & \left(p_{xv}^*\right)_{k} \Big/ \sqrt{\left(p_x^* \; p_v^*\right)_{k-1/k-1}} \\ \\ \left(p_{xv}^*\right)_{k} \Big/ \sqrt{\left(p_x^* \; p_v^*\right)_{k-1/k-1}} & \left(p_v^*\right)_{k} \Big/ \left(p_v^*\right)_{k-1/k-1} \end{bmatrix} \tag{121}$$

$$\tau = \tau^* \sqrt{\left(\frac{p_v^*}{p_x^*}\right)_{k-1/k-1}} \tag{122}$$

$$\Phi = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix} \tag{123}$$

$$H = [1 \quad 0] \tag{124}$$

$$r_x = r_x^*/\left(p_x^*\right)_{k-1/k-1} \tag{125}$$

$$q_x = q_x^*/\left(p_x^*\right)_{k-1/k-1} \tag{126}$$

$$q_v = q_v^*/\left(p_v^*\right)_{k-1/k-1} \quad . \tag{127}$$

$(p_x)_{k/k-1}$, the predicted next-stage position covariance, and $(p_x)_{k/k}$, the corrected next-stage position covariance, are plotted in Fig. 15. Using the following equations [see Eqs. (89)-(99)],

$$\left(p_x\right)_{k/k-1} = 1 + a\tau + \tau^2 \tag{128}$$

$$\left(p_x\right)_{k/k} = \frac{r_x\left(p_x\right)_{k/k-1}}{r_x + \left(p_x\right)_{k/k-1}} = \frac{1}{\frac{1}{r_x} + \frac{1}{\left(p_x\right)_{k/k-1}}} \quad . \tag{129}$$

The parameters are, besides $(p_x)_{k/k-1}$ and τ, the position measurement noise, r_x, and the time rate of change of the predicted covariance, denoted by a:

$$a = 2\left(p_{xv}\right)_{k-1/k-1} \; ; \tag{130}$$

if the plant noise q_x is independent of the sampling period τ.

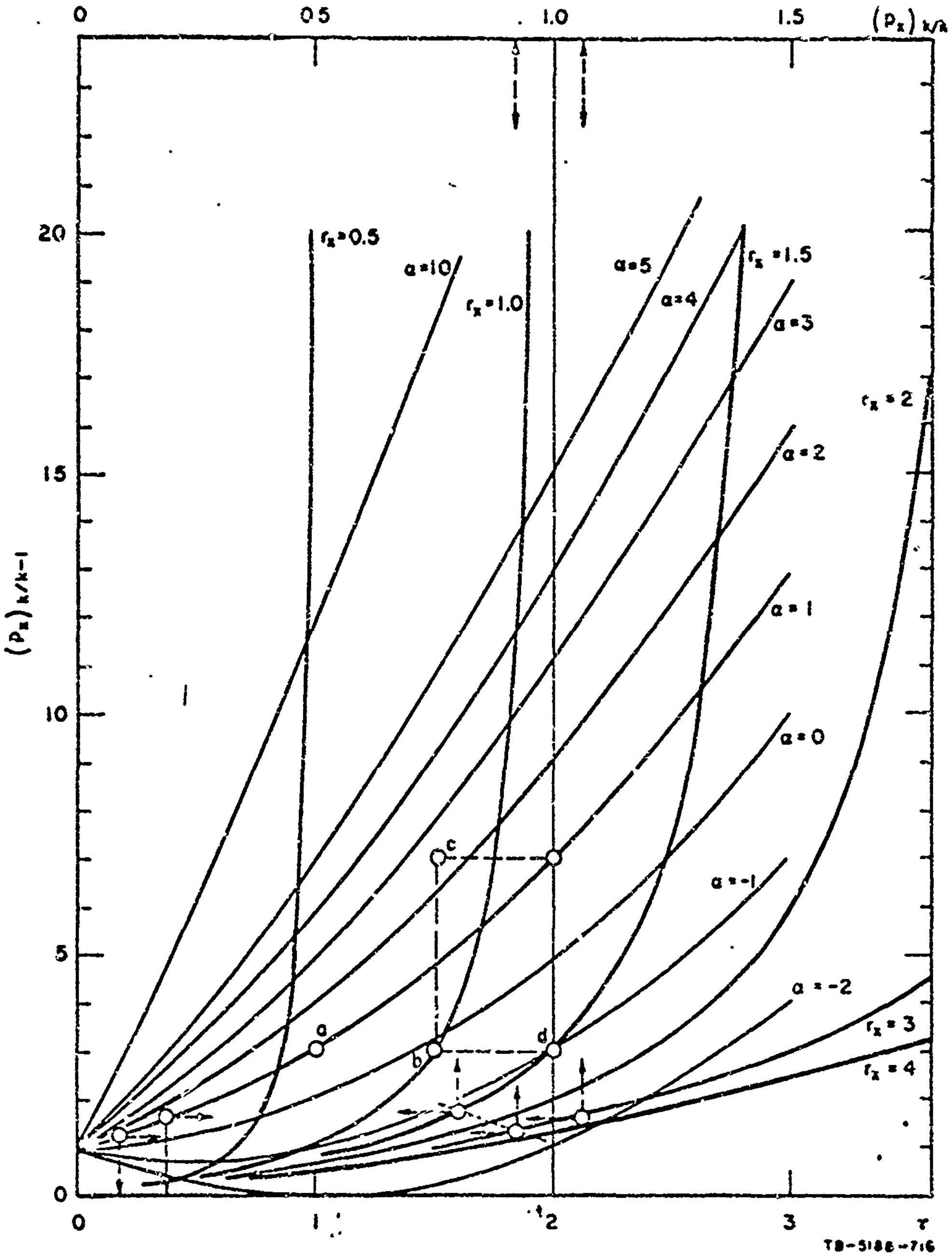

FIG. 15 EFFECT OF SAMPLING PERIOD FOR $1/s^2$ PLANT WITH POSITION MEASUREMENT

The value of a is

$$a = 2\left(p_{xv}\right)_{k-1/k-1} + \alpha \tag{131}$$

if the plant noise q_x depends on the sampling period, e.g., $q_x = \alpha T$.

$$r_x = \text{constant, if position data is not smoothed.} \tag{132}$$

$$r_x = r/\tau\text{, if position data is perfectly smoothed} \tag{133}$$

(r is a position noise parameter).

Figure 15 is organized as follows: the lower horizontal axis is for τ, the vertical axis for $(p_x)_{k/k-1}$, and the upper horizontal axis for $(p_x)_{k/k}$. Therefore, if the problem is to find $(p_x)_{k/k}$ for a given set of τ, r_x, and a, we start from the lower horizontal axis, where the given τ is located. We then traverse up vertically until the appropriate r_x curve is reached. $(p_x)_{k/k-1}$ can now be read from the vertical axis. To continue, we traverse horizontally to the curve with appropriate value of a. The upper horizontal axis now gives the answer on $(p_x)_{k/k}$. For other types of problems, the stated procedure may be modified easily.

Figures 16 and 17 show the relationship between the sensitivity, $(\Delta p_x/p_x)_{k/k}/(\Delta\tau/\tau)$, and τ for a = 0 and 1. The equation for these figures is easily derived from Eqs. (128) and (129):

$$\left(\frac{\Delta p_x}{p_x}\right)_{k/k}\Bigg/\left(\frac{\Delta\tau}{\tau}\right) = \frac{\tau\ (a + 2\tau)}{\left(\dfrac{1 + a\tau + \tau^2}{r_x} + 1\right)(1 + a\tau + \tau^2)} \quad . \tag{134}$$

Note that this sensitivity is dimensionless, relating percentage change in $(p_x^*)_{k/k}$ to percentage change in τ^*:

$$\left(\frac{\Delta p_x^*}{p_x^*}\right)_{k/k}\Bigg/\left(\frac{\Delta\tau^*}{\tau^*}\right) = \left(\frac{\Delta p_x}{p_x}\right)_{k/k}\Bigg/\left(\frac{\Delta\tau}{\tau}\right) \quad . \tag{135}$$

Two examples are given below illustrating the use of Fig. 15.

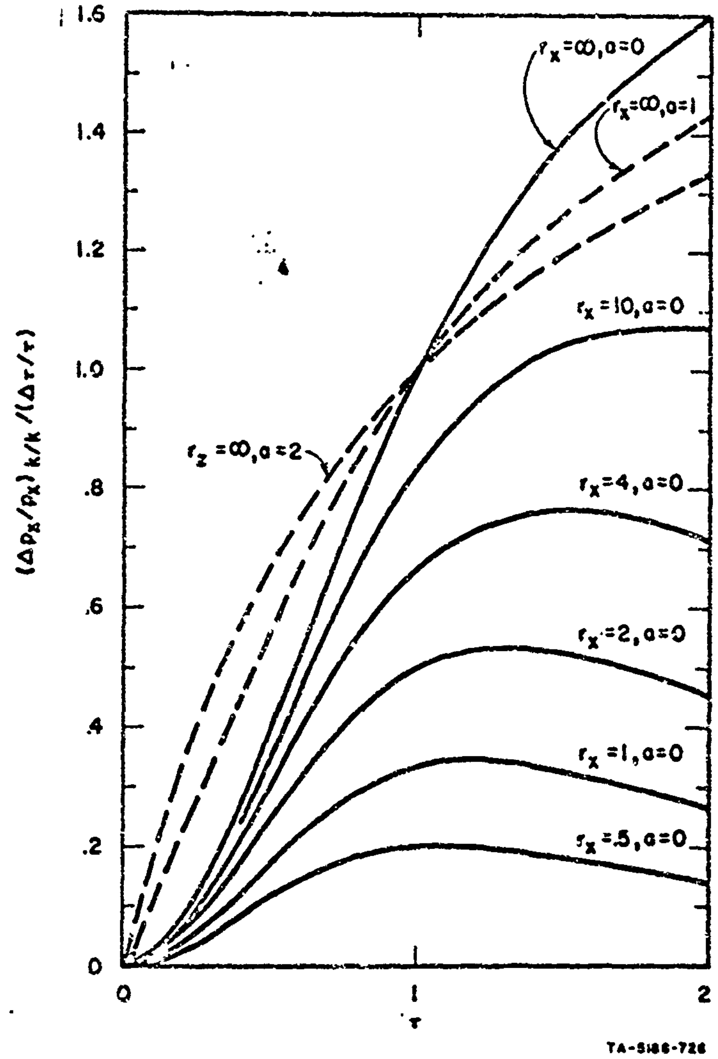

FIG. 16 SAMPLING-PERIOD SENSITIVITY OF ONE-STAGE KALMAN FILTERING, $a = 0$

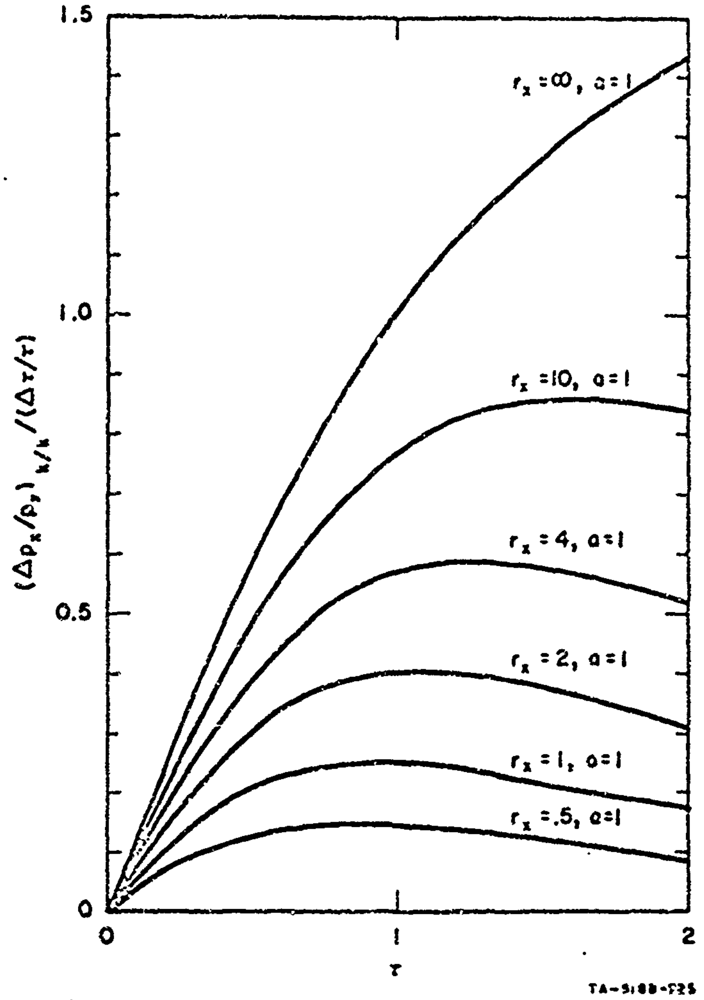

FIG. 17 SAMPLING-PERIOD SENSITIVITY OF ONE-STAGE KALMAN FILTERING, $a = 1$

Example 1:

Change in sampling period

$$p^*_{k-1/k-1} = \begin{bmatrix} 3,422 & 5,380 \\ 5,380 & 11,973 \end{bmatrix}$$

$$r^*_x = 10^4 \text{ ft}^2 \tag{136}$$

$$q^*_x = 0.259 \times \tau^* \text{ ft}^2 \quad .$$

We wish to find and compare $(p^*_x)_{k/k}$ for sampling periods, τ^*, of 0.1 and 0.2 sec. According to Eqs. (118)-(127), the normalized quantities are

$$P_{k-1/k-1} = \begin{bmatrix} 1 & \frac{5380}{6400} \\ \frac{5380}{6400} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0.840 \\ 0.840 & 1 \end{bmatrix}$$

$$\begin{aligned} \tau = \tau^* \times 1.87 &= 0.187 \quad \text{if} \quad \tau^* = 0.1 \\ &= 0.374 \quad \text{if} \quad \tau^* = 0.2 \end{aligned} \tag{137}$$

$$r_x = 10^4/3422 = 2.92$$

$$q_x = \frac{0.259 \times \tau^*}{3422} = \frac{0.259}{3422 \times 1.87}\tau = 0.0000405\tau \quad .$$

Therefore,

$$\alpha = 0.0000405$$

and

$$\begin{aligned} a &= 2x \left(p_{xv}\right)_{k-1/k-1} + \alpha \\ &= 2x\ 0.840 + 0.0000405 = 1.68 \quad . \end{aligned} \tag{138}$$

To obtain values of $(p_x)_{k/k}$ for τ of 0.187 and 0.374, the dotted lines in Fig. 15 are traversed from τ axis to $(p_x)_{k/k}$ axis. The results

$$\begin{aligned} \left(p_x\right)_{k/k} &= 0.92 \quad \text{for} \quad \tau = 0.187 \\ \left(p_x\right)_{k/k} &= 1.06 \quad \text{for} \quad \tau = 0.374 \quad . \end{aligned} \tag{139}$$

The conclusions are therefore thus: If a sampling period of 0.1 sec ($\tau^* = 0.1$) is used, the one-stage position covariance reduction is 0.92, from 3,422 to 3,150. If the sampling period is increased to 0.2 sec, the position covariance increases by a factor of 1.06, from 3,422 to 3,630. Furthermore, this increase is due mainly to velocity errors rather than to plant noise; this may be seen in Eq. (138), where $(p_{xv})_{k-1/k-1}$ is due to velocity errors, and α is due to plant noise.

These values are verified by Fig. 7 at 1 sec.

Example 2:

Sampling-period change with smoothed measurement

Assume all quantities have values identical to the last example with the exception of r_x:

$$r_x = \frac{0.187 \times 2.92}{\tau} \quad . \tag{140}$$

For $\tau^* = 0.1$ and 0.2, r_x is 2.92 and 1.46, respectively. The first case τ^* of 0.1 is read from Fig. 15 exactly as in the last example, which gives $(p_x)_{k/k} = 0.92$. For $\tau^* = 0.2$, we have to use the $r_x = 1.46$ curve (approximated as shown in dotted line by $r_x = 1.50$), obtaining $(p_x)_{k/k} = 0.8$.

Based on the values, we have the conclusion that for larger position covariance reduction using smoothed data, Eq. (140), one should use a sampling period of 0.2 sec compared to 0.1 sec. If this seems unreasonable according to engineering intuition, one should refer to the discussions in Example 2 of Sec. III-E.

For τ^* between 0.1 and 0.2 (τ between 0.187 and 0.374) one may interpolate using the -.- line shown in Fig. 15. In the smoothed data case then, this line takes the place of constant r_x lines of the non-smoothed data case.

Perhaps the greatest utility of Fig. 15 is in obtaining rough and quick estimation of the filter characteristics at the operating points. For example, when r_x is small and when $(p_x)_{k/k-1}$ is large, changes in sampling period do not affect $(p_x)_{k/k}$ very much--see, as specific values, $r_x = 0.5$, $a = 1$, and $\tau = 1$ in Fig. 15. Another example is the trade-off example shown in the following.

Example 3:

Trade-offs

Let the filter operation be represented by the two points a and b in Fig. 15, representing $\tau = 1$, $a = 1$, and $r_x = 1$. $(p_x)_{k/k}$ is 0.75. Suppose in order to reduce computational load we wish to change τ from 1 to 2; we have a choice of either improving model accuracy or improving measurement accuracy to maintain $(p_x)_{k/k}$ at 0.75.

First let us change model accuracy while r_x is held fixed. As τ is changed to 2, $a = -1$ is required (point d) to maintain $(p_x)_{k/k}$ at 0.75. If $(p_x)_{k-1/k-1}$ is positive, a negative a is impossible because by Eq. (131) α is a positive number. This leaves us the only choice of reducing the measurement noise. This is represented by the point c. The required r_x is about 0.8. The measurement noise covariance therefore needs to be reduced by $(1 - 0.8)/1 = 20$ percent.

V CONCLUSIONS AND FUTURE WORK

We have given techniques for sensitivity analysis of the Kalman filter with respect to simultaneous variations in measurement noise, plant noise, dynamic model, sampling period, and filter gain. These analytical techniques will greatly aid the design and evaluation of Kalman filters and other types of filters. Two basic assumptions were used:

(1) There are nominal quantities about which variations may be taken.

(2) The estimation-error covariances are the filter performance measures.

Future work shall be the application of the techniques to the problems described in Sec. II-G.

REFERENCES

1. R. E. Larson, R. M. Dressler, and R. S. Ratner, "Application of the Extended Kalman Filter to Ballistic Trajectory Estimation," Final Report, SRI Project 5188-103 for NIKE-X Project Office (January 1967).

2. L. Meier, J. Peschon, B. Ho, R. Larson, and R. M. Dressler, "Design of Guidance and Control Systems for Optimal Utilization of Information," Final Report, SRI Project 5067 for NASA Ames Research Center (May 1967).

3. B. Friedland, "On the Effect of Incorrect Gain in Kalman Filter," IEEE Trans. on Automatic Control, p. 610 (October 1967).

4. T. Nishimura, "On the a priori Information in Sequential Estimation Problems," IEEE Trans. on Automatic Control, Vol. AC-11, pp. 197-204 (April 1966).

5. T. Nishimura, "Error Bounds of Continuous Kalman Filter and the Application to Orbit Determination Problems," IEEE Trans. on Automatic Control, Vol. AC-12, No. 3, pp. 268-275 (June 1967).

6. H. Heffes, "The Effect of Erroneous Models on the Kalman Filter Response," IEEE Trans. on Automatic Control, pp. 541-543 (July 1966).

7. S. L. Fagin, "Recursive Linear Regression Theory, Optimal Filter Theory, and Error Analysis of Optimal Systems," 1964 IEEE International Conv. Record, Part I, pp. 216-240.

8. R. E. Larson, R. M. Dressler, R. S. Ratner, and B. L. Ho, "Applications of the Extended Kalman Filter to BRV and MRV Trajectory Estimation," Interim Report, SRI Project 5188-203 for NIKE-X Project Office (December 1967), CONFIDENTIAL.

9. R. H. Battin, Astronautical Guidance (McGraw-Hill Book Co., 1964).

APPENDIX A

ACTUAL COVARIANCES IN EXTENDED KALMAN FILTER

The extended Kalman filter is becoming an important technique estimation of nonlinear dynamical system. The purpose of this appendix is to give expressions for the expectations and of the estimation-error covariances using the extended Kalman filter, and to reconcile them with the covariance equation of the filter. The discrete time case shall be considered.

The system is described by

$$\begin{aligned} x_{k+1} &= f_k(x_k, w_k) \\ z_k &= h_k(x_k, v_k) \quad . \end{aligned} \tag{A.1}$$

We note that the dynamic equations of Sec. II, Eq. (1) are a special case of the above equations.

We will analyze the following filter:

$$\begin{aligned} \hat{x}_{k/k-1} &= f_{k-1}(\hat{x}_{k-1/k-1}, 0) \\ \hat{z}_{k/k-1} &= h_k(\hat{x}_{k/k-1}, 0) \\ \hat{x}_{k/k} &= \hat{x}_{k/k-1} + W_k(z_k - \hat{z}_{k/k-1}) \quad , \end{aligned} \tag{A.2}$$

where f and h are nonlinear functions with sufficient smootheness for our later derivations, x is the state to be estimated, z the measurement, and w and v the plant and measurement noises. The noise are assumed to be zero mean and uncorrelated both to each other and in time. W is the filter gain; the super-hat symbol denotes the estimated quantities.

The estimation errors are defined as

$$\tilde{x}_{k/k-1} = x_k - \hat{x}_{k/k-1}$$
$$\tilde{x}_{k/k} = x_k - \hat{x}_{k/k} \quad . \tag{A.3}$$

Assume that at the $(k - 1)^{th}$ stage, we have

$$E[\tilde{x}_{k-1/k-1}] = \bar{\tilde{x}}_{k-1/k-1}$$
$$E\left[(\tilde{x}_{k-1/k-1} - \bar{\tilde{x}}_{k-1/k-1})(\tilde{x} - \bar{\tilde{x}}_{k-1/k-1})^T\right] = P_{k-1/k-1} \quad . \tag{A.4}$$

For the prediction part of the filtering, we have

$$\begin{aligned}\tilde{x}_{k/k-1} &= x_k - \hat{x}_{k/k-1} \\ &= f_{k-1}(x_{k-1}, w_{k-1}) - f_{k-1}(\hat{x}_{k-1/k-1}, 0) \quad .\end{aligned}$$

It is assumed that $f(x, w)$ may be approximated by a two-term Taylor series about $f(\hat{x}, 0)$. Dropping the subscript $k - 1$ on f_{k-1} for simplicity, we obtain

$$\begin{aligned}\tilde{x}_{k/k-1} = &\Big[f(\hat{x}_{k-1/k-1}, 0) + f_x \cdot (x_{k-1} - \hat{x}_{k-1/k-1}) + f_w \circ w_{k-1} \\ &+ \frac{1}{2} f_{xx} \circ (x_{k-1} - \hat{x}_{k-1/k-1})(x_{k-1} - \hat{x}_{k-1/k-1})^T \\ &+ \frac{1}{2} f_{ww} \circ w_{k-1} w_{k-1}^T + f_{xw} \circ (x_{k-1} - \hat{x}_{k-1/k-1}) w_{k-1}^T\Big] \\ &- f(\hat{x}_{k-1/k-1}, 0)\end{aligned} \tag{A.5}$$

where the following notation is used. Let f be an n-vector function of an n-vector x and a p-vector y, and let B be an $n \times p$ matrix. $f_{xy} \circ B$ will denote an n-vector, the kth component of which is

$$\left(f_{xy} \circ B\right)_k \triangleq \sum_i \sum_j \frac{\partial^2 f_k}{\partial x_i \, \partial y_j} B_{ij} \quad . \tag{A.6}$$

The special case of $y = x$ and $B = \delta x \; \delta x^T$ is the following familiar term in the Taylor expansion

$$\left(f_{xx} \circ \delta x \; \delta x^T\right)_k = \sum_i \sum_j \frac{\partial^2 f_k}{\partial x_i \, \partial y_i} \delta x_i \; \delta x_j \quad . \tag{A.7}$$

Note that

$$f_{xy} \circ (A + B) = f_{xy} \circ A + f_{xy} \circ B \quad . \tag{A.8}$$

Therefore,

$$\begin{aligned} \tilde{x}_{k/k-1} &= f_x \cdot \tilde{x}_{k-1/k-1} + f_w \cdot w_{k-1} + \frac{1}{2} f_{xx} \circ \tilde{x}_{k-1/k-1} \tilde{x}^T_{k-1/k-1} \\ &+ \frac{1}{2} f_{ww} \circ w_{k-1} w^T_{k-1} + f_{xw} \circ \tilde{x}_{k-1/k-1} w^T_{k-1} \quad . \end{aligned} \tag{A.9}$$

Therefore, upon using Eq. (A.5), and since w_{k-1} and $\tilde{x}_{k-1/k-1}$ are uncorrelated and $E[w_{k-1}] = 0$, the expected value of $x_{k/k-1}$ is

$$\begin{aligned} \bar{\tilde{x}}_{k/k-1} = E[\tilde{x}_{k/k-1}] &= f_x \cdot \bar{\tilde{x}}_{k-1/k-1} \\ &+ \frac{1}{2} f_{xx} \circ \left[P_{k-1/k-1} + \bar{\tilde{x}}_{k-1/k-1} \bar{\tilde{x}}^T_{k-1/k-1}\right] \\ &+ \frac{1}{2} f_{ww} \circ Q_{k-1} \end{aligned} \tag{A.10}$$

$$\begin{aligned} (\tilde{x} - \bar{\tilde{x}})_{k/k-1} &= f_x \cdot (\tilde{x}_{k-1/k-1} - \bar{\tilde{x}}_{k-1/k-1}) + f_w \cdot w_{k-1} \\ &+ \frac{1}{2} f_{xx} \circ \left[(\tilde{x} \, \tilde{x}^T)_{k-1/k-1} - (\bar{\tilde{x}} \, \bar{\tilde{x}}^T)_{k-1/k-1} - P_{k-1/k-1}\right] \\ &+ \frac{1}{2} f_{ww} \circ \left[(w \, w^T)_{k-1} - Q_{k-1}\right] + f_{xw} \circ \tilde{x}_{k-1/k-1} w^T_{k-1} \quad . \end{aligned} \tag{A.11}$$

Now we compute the covariance

$$P_{k/k-1} = E\left[(\tilde{x}_{k/k-1} - \bar{\tilde{x}}_{k/k-1})(\tilde{x}_{k/k-1} - \bar{\tilde{x}}_{k/k-1})^T\right]$$

$$P_{k/k-1} \cong f_x\, E\left[(\tilde{x}_{k-1/k-1} - \bar{\tilde{x}}_{k-1/k-1})(\tilde{x}_{k-1/k-1} - \bar{\tilde{x}}_{k-1/k-1})^T\right] f_x^T$$

$$+ f_w\, E\left[w_{k-1}\, w_{k-1}^T\right] f_w^T$$

$$= f_x\, P_{k-1/k-1}\, f_x^T + f_w\, Q_{k-1}\, f_w^T \quad , \tag{A.12}$$

where all terms involving third or higher orders of $\tilde{x}_{k-1/k-1}$ and w_{k-1} are assumed to be negligible compared to $P_{k/k-1}$. This is a valid assumption when either $\tilde{x}_{k-1/k-1}$ and w_{k-1} are small or $\left(\partial^2 f_i\right)/(\partial x_j\, \partial x_k)$ and $\left(\partial^2 f_i\right)/(\partial x_j\, \partial w_j)$ are small.

For the updating part of the filter, we have (dropping the subscript k on W and h)

$$\tilde{x}_{k/k} = x_k - \hat{x}_{k/k}$$

$$= x_k - [\hat{x}_{k/k-1} + W(z_k - \hat{z}_{k/k-1})]$$

$$= x_k - \hat{x}_{k/k-1} - Wh(x_k, v_k) + Wh(\hat{x}_{k/k-1}, 0) \quad .$$

Using a second-order approximation of $h(x_k, v_k)$, we have

$$\tilde{x}_{k/k} = \tilde{x}_{k/k-1} - W\Big[h(\hat{x}_{k/k-1}, 0) + h_x(x_k - \hat{x}_{k/k-1}) + h_v\, v_k$$

$$+ \frac{1}{2}\, h_{xx} \circ (x_k - \hat{x}_{k/k-1})(x_k - \hat{x}_{k/k-1})^T + \frac{1}{2}\, h_{vv} \circ v_k\, v_k^T$$

$$+ h_{xv} \circ (x_k - \hat{x}_{k/k-1})\, v_k^T\Big] + Wh(\hat{x}_{k/k-1}, 0) \quad .$$

Therefore,

$$\tilde{x}_{k/k} = (I - Wh_x)\ \tilde{x}_{k/k-1} - Wh_v\ v_k - \frac{1}{2} Wh_{xx} \circ \tilde{x}_{k/k-1}\ \tilde{x}^T_{k/k-1}$$

$$- \frac{1}{2} Wh_{vv} \circ v_k\ v_k^T - Wh_{xv} \circ \tilde{x}_{k/k-1}\ v_k^T \quad . \qquad (A.13)$$

Using the last expression, the expectation and covariance of $\tilde{x}_{k/k}$ may be computed, upon using smallness assumptions similar to the prediction part for $P_{k/k}$:

$$\bar{\tilde{x}}_{k/k} = E[\tilde{x}_{k/k}] = (I - Wh_x)\ \bar{\tilde{x}}_{k/k-1}$$

$$- \frac{1}{2} Wh_{xx} \circ \left(P_{k/k-1} + \bar{\tilde{x}}_{k/k-1}\ \bar{\tilde{x}}^T_{k/k-1}\right) - \frac{1}{2} Wh_{vv} \circ R_k \qquad (A.14)$$

$$\tilde{x}_{k/k} - \bar{\tilde{x}}_{k/k} = (I - Wh_x)(\tilde{x} - \bar{\tilde{x}}) - Wh_v\ v_k$$

$$- \frac{1}{2} Wh_{xx} \circ \left[(\tilde{x}\ \tilde{x}^T)_{k/k-1} - (\bar{\tilde{x}}\ \bar{\tilde{x}}^T)_{k/k-1} - P_{k/k-1}\right]$$

$$- \frac{1}{2} Wh_{vv} \circ \left(v_k\ v_k^T - R_k\right) - \frac{1}{2} Wh_{xv} \circ \tilde{x}_{k/k-1}\ v_k^T \qquad (A.15)$$

$$P_{k/k} = E\left[(\tilde{x}_{k/k} - \bar{\tilde{x}}_{k/k})(\tilde{x}_{k/k} - \bar{\tilde{x}}_{k/k})^T\right]$$

$$\cong (I - Wh_x)\ P_{k/k-1}\ (I - Wh_x) + Wh_v\ k_k\ h_v^T\ W^T \quad . \qquad (A.16)$$

A comparison of the actual covariance expressions, Eqs. (A.12) and (A.16), with the computed covariance equations of the extended Kalman filter shows that they are identical. The computed covariance equations therefore give the actual covariances approximated to the second order. However, the estimation-error means given in Eqs. (A.10) and (A.14) are no longer zero, as is true in a linear system.

By modifying the extended Kalman filter equations of Eq. (A.2), the estimation-error means may be made zero. The optimality of the

modified filter will not be discussed here. The modifications consists of adding second-order terms as follows:

$$\hat{x}_{k/k-1} = f(\hat{x}_{k-1/k-1}, 0) - \frac{1}{2} f_{xx} \circ P_{k-1/k-1} - \frac{1}{2} f_{ww} \circ Q_{k-1}$$

$$\hat{z}_{k/k-1} = h(\hat{x}_{k/k-1}, 0)$$

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + \frac{1}{2} Wh_{xx} \circ P_{k/k-1} + \frac{1}{2} Wh_{vv} \circ R_k$$

$$\hat{x}_{0/-1} = \bar{\bar{x}}_{0/-1} \qquad \text{(A.17)}$$

For this filter, $\bar{\bar{x}}_{k/k-1} = 0$ and $\tilde{x}_{k/k} = 0$, with the actual covariances given by Eqs. (A.12) and (A.16).

APPENDIX B

THE VARIATIONAL EQUATION OF THE COVARIANCE EQUATIONS

In this appendix, the change in $P_{k+1/k+1}$ as a result of changes in $P_{k/k}$, Φ_k, Q_k, R_{k+1} will be studied. For simplicity of notation, all subscripts will be dropped on Φ_k, Q_k, R_{k+1}, and W_{k+1}. P_{k+1} denotes $P_{k+1/k+1}$, P denotes $P_{k+1/k}$, and P_k denotes $P_{k/k}$.

The prediction-error covariance matrix when P_k, Φ, and Q are changed to $P_k + \Delta P_k$, $\Phi + \Delta\Phi$, and $Q + \Delta Q$ is

$$P + \Delta P = (\Phi + \Delta\Phi)\ (P_k + \Delta P_k)\ (\Phi + \Delta\Phi)^T + (Q + \Delta Q)$$

$$= \left(\Phi P_k \Phi^T + Q\right) + \Delta\Phi P_k \Phi^T + \Phi P_k \Delta\Phi^T + \Delta\Phi P_k \Delta\Phi^T$$

$$+ (\Phi + \Delta\Phi)\ \Delta P_k\ (\Phi + \Delta\Phi)^T + \Delta Q \quad . \tag{B.1}$$

Since $P = \Phi P_k \Phi^T + Q$, we have

$$\Delta P = \left(\Phi \Delta P_k \Phi^T + \Delta Q\right) + \left(\Delta\Phi P_k \Phi^T + \Phi P_k \Delta\Phi^T\right)$$

$$+ \left(\Delta\Phi P_k \Delta\Phi^T + \Delta\Phi \Delta P_k \Phi^T + \Phi \Delta P_k \Delta\Phi^T\right)$$

$$+ \left(\Delta\Phi \Delta P_k \Delta\Phi^T\right) \quad . \tag{B.2}$$

If $\Delta\Phi = 0$, only the first bracketed term is nonzero. For nonzero $\Delta\Phi$, ΔP is given to the first order by the first two bracketed terms of Eq. (B.2).

The updated error covariance P_{k+1} changes because of ΔP, ΔR, and ΔW. For the general case with no assumptions on ΔW, we have

$$P_{k+1} + \Delta P_{k+1} = [I - (W + \Delta W)H]\ [P + \Delta P]\ [I - (W + \Delta W)H]^T$$

$$+ (W + \Delta W)\ (R + \Delta R)\ (W + \Delta W)^T \quad . \tag{B.3}$$

Upon subtracting the expression

$$P_{k+1} = [I - WH]\, P[I - WH]^T + WRW^T. \qquad (B.4)$$

and gathering terms, we have

$$\Delta P_{k+1} = \{(I - WH)\ \Delta P(I - WH)^T + W\Delta RW^T\} + A_{k+1} + B_{k+1} \quad , \qquad (B.5)$$

where

$$A_{k+1} = \{[-(I - WH)\ PH^T + WR]\ \Delta W^T + \Delta W[-(I - WH)\ PH^T + WR]^T\}$$

$$B_{k+1} = \{\Delta W[HPH^T + R]\ \Delta W^T + [-(I - WH)\ \Delta PH^T + W\Delta R]\ \Delta W^T$$

$$+ \Delta W[-(I - WH)\ \Delta PH^T + W\Delta R]^T\}$$

$$+ \{\Delta W[H\Delta PH^T + \Delta R]\ \Delta W^T\} \quad .$$

The first bracketed term and A_{k+1} contain the first-order terms; the B_{k+1} is of second and third order.

If the gain is fixed, $\Delta W \simeq 0$, then only the first bracketed term is nonzero.

If the nominal gain W is optimal (before changes in P and R), then A_{k+1} is zero, regardless of ΔW, because

$$W^o = PH^T(HPH + R)^{-1} \qquad (B.6)$$

or

$$[-(I - W^oH)\ PH^T + W^oR] = 0 \quad , \qquad (B.7)$$

but B_{k+1} assumes some nonzero value depending on ΔW.

We have, in fact, obtained a verification that W^o, as shown in Eq. (B.6), is optimum in the sense that it causes the smallest estimation-error covariance. The verification consists of showing that any variation from W^o will result in higher covariance $P_{k+1/k+1}$. Accordingly, we set $\Delta P = 0$, $\Delta R = 0$ in Eq. (B.5), obtaining the following B_{k+1} that is positive definite for all ΔK--$\Delta W[HPH^T + R]\,\Delta W^T$. Therefore, if W of Eq. (B.6) is used, then A_{k+1} is zero, and $\Delta P_{k+1} = B_{k+1}$ is positive definite. This verifies the optimality of W in Eq. (B.6).

Let us now investigate the case in which W is arbitrary but ΔW is such that $(W + \Delta W)$ is optimum under parameter variation $P + \Delta P$ and $R + \Delta R$. We shall denote this ΔW by ΔW^o. By Eq. (B.6),

$$W + \Delta W^o = (P + \Delta P)H^T\,[H(P + \Delta P)H^T + (R + \Delta R)]^{-1} \quad , \tag{B.8}$$

which may be rearranged as

$$\Delta W^o = \pi_2\,\pi_1^{-1} \tag{B.9}$$

where

$$\pi_1 = H(P + \Delta P)H^T + R + \Delta R$$

$$\pi_2 = (I - WH)\,(P + \Delta P)H^T - W(R + \Delta R) \quad . \tag{B.10}$$

Now A_{k+1} and B_{k+1} may be expressed as

$$A_{k+1} + B_{k+1} = (\Delta W)\,\pi_1(\Delta W)^T - \pi_2(\Delta W)^T - (\Delta W)\,\pi_2 \quad . \tag{B.11}$$

Using the optimum ΔW^o of Eq. (B.9), we have

$$A_{k+1} + B_{k+1} = \pi_2\,\pi_1^{-1}\,\pi_2^T - \pi_2\,\pi_1^{-1}\,\pi_2^T - \pi_2\,\pi_1^{-1}\,\pi_2^T$$

$$= -\,\pi_2\,\pi_1^{-1}\,\pi_2^T \quad . \tag{B.12}$$

Therefore, if ΔW^o is used regardless of the optimality of the nominal W, $A_{k+1} + B_{k+1}$ is the negative definite matrix of Eq. (B.12).

Of particular interest is the case when W and W + ΔW are both optimal:

$$\Delta P_{k+1} = (I - W^o H)\ \Delta P(I - W^o H)^T + W^o\ \Delta R W^{oT}$$
$$- [-\Delta P H^T + W^o(H\Delta P H^T + \Delta R)]\ [H(P + \Delta P)\ H^T + R + \Delta R]^{-1}$$
$$[-\Delta P H^T + W(H\Delta P H^T + \Delta R)]^T \quad , \qquad (B.13)$$

and when W is fixed, ($\Delta W = 0$), regardless of the optimality of W:

$$\Delta P_{k+1} = (I - WH)\ \Delta P(I - WH)^T + W\Delta R W^T \quad . \qquad (B.14)$$

APPENDIX C

NORMALIZATION OF THE COVARIANCE EQUATION

The recursive covariance equations are

$$P^*_{k/k-1} = \Phi^* P^*_{k-1/k-1} \Phi^{*T} + Q^*$$

$$P^*_{k/k} = P^*_{k/k-1} - P^*_{k/k-1} H^{*T} \left(H^* P^*_{k/k-1} H^{*T} + R^*_k\right)^{-1} H^*_k P^*_{k/k-1} \quad , \tag{C.1}$$

where

$$P^*_{k-1/k-1} = \begin{bmatrix} \left(P^*_x\right)_{k-1/k-1} & \left(P^*_{xv}\right)_{k-1/k-1} \\ \left(P^*_{xv}\right)_{k-1/k-1} & \left(P^*_v\right)_{k-1/k-1} \end{bmatrix} \tag{C.2}$$

$$\Phi^* = \begin{bmatrix} 1 & T^* \\ 0 & 1 \end{bmatrix} \tag{C.3}$$

$$H^* = [h_1 \quad h_2] \quad . \tag{C.4}$$

Multiplying both sides of Eq. (C.1) by the normalization matrix

$$N^{-1/2} = \begin{bmatrix} \dfrac{1}{\sqrt{\left(P^*_x\right)_{k-1/k-1}}} & 0 \\ 0 & \dfrac{1}{\sqrt{\left(P^*_v\right)_{k-1/k-1}}} \end{bmatrix} \quad , \tag{C.5}$$

we obtain the following normalized equations:

$$
P_{k/k-1} = \Phi\, P_{k-1/k-1}\, \Phi^T + Q
$$

$$
P_{k/k} = P_{k/k-1} - P_{k/k-1}\, H^T \left(H\, P_{k/k-1}\, H^T + R_k\right)^{-1} H\, P_{k/k-1} \quad , \tag{C.6}
$$

where

$$
P_{k-1/k-1} = N^{-1/2}\, P^*_{k-1/k-1}\, N^{-1/2} = \begin{bmatrix} 1 & \dfrac{\left(p_{xv}\right)_{k-1/k-1}}{\left(\sqrt{p_x\, p_v}\right)_{k-1/k-1}} \\ \dfrac{\left(p_{xv}\right)_{k-1/k-1}}{\left(\sqrt{p_x\, p_v}\right)_{k-1/k-1}} & 1 \end{bmatrix} \tag{C.7}
$$

$$
P_{k/k-1} = N^{-1/2}\, P^*\, N^{-1/2} = \begin{bmatrix} \dfrac{\left(p_x\right)_{k/k-1}}{\left(p_x\right)_{k-1/k-1}} & \dfrac{\left(p_{xv}\right)_{k/k-1}}{\left(\sqrt{p_x\, p_v}\right)_{k-1/k-1}} \\ \dfrac{\left(p_{xv}\right)_{k/k-1}}{\left(\sqrt{p_x\, p_v}\right)_{k-1/k-1}} & \dfrac{\left(p_v\right)_{k/k-1}}{\left(p_v\right)_{k-1/k-1}} \end{bmatrix} \tag{C.8}
$$

$$
P_{k/k} = N^{-1/2}\, P^*\, N^{-1/2} = \begin{bmatrix} \dfrac{\left(p_x\right)_{k/k}}{\left(p_x\right)_{k-1/k-1}} & \dfrac{\left(p_{xv}\right)_{k/k}}{\left(\sqrt{p_x\, p_v}\right)_{k-1/k-1}} \\ \dfrac{\left(p_{xv}\right)_{k/k}}{\left(\sqrt{p_x\, p_v}\right)_{k-1/k-1}} & \dfrac{\left(p_v\right)_{k/k}}{\left(p_v\right)_{k-1/k-1}} \end{bmatrix} \tag{C.9}
$$

$$
\Phi = N^{-1/2}\, \Phi^*\, N^{1/2} = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix} \quad , \qquad \tau = (\tau^*) \left(\sqrt{\frac{p_v^*}{p_x^*}}\right)_{k-1/k-1} \quad . \tag{C.10}
$$

The measurement vector H^* can be normalized to [1 0] by transforming H^* and R^* using some nonsingular matrix A in the following manner. This will leave $P_{k/k}$ unchanged; verification is by substituting Eqs. (C.11), (C.12) and Eqs. (C.7)-(C.10) into Eq. (C.1).

$$
H = [1 \quad 0]\, N^{-1/2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} A^{-1}\, H^*\, N^{1/2} = [1 \quad 0] \tag{C.11}
$$

$$R_k = [1 \quad 0]\, N^{-1/2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} A^{-1}\, R_k^*\, A^{T-1}\, [1 \quad 0]\, N^{-1/2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad . \quad \text{(C.12)}$$

Similar equations exist for H^* that are matrices rather than vectors. Special cases of H^* are:

$$\text{if} \quad H^* = [1 \quad 0] \ , \quad \text{then} \quad H = [1 \quad 0] \quad \text{and} \quad R = \frac{R_x}{\left(p_x\right)_{k-1/k-1}} \quad ; \quad \text{(C.13)}$$

$$\text{if} \quad H^* = [0 \quad 1] \ , \quad \text{then} \quad H = [0 \quad 1] \quad \text{and} \quad R = \frac{R_v}{\left(p_v\right)_{k-1/k-1}} \quad ; \quad \text{(C.14)}$$

$$\text{if} \quad H^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \ , \quad \text{then} \quad H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad R = \begin{bmatrix} \dfrac{R_x}{\left(p_x\right)_{k-1/k-1}} & 0 \\ 0 & \dfrac{R_v}{\left(p_v\right)_{k-1/k-1}} \end{bmatrix}$$

(C.15)

$$Q = N^{-1/2}\, Q^*\, N^{-1/2} = \begin{bmatrix} \dfrac{q_x}{\left(p_x\right)_{k-1/k-1}} & \dfrac{q_{xv}}{\left(\sqrt{p_x\, p_v}\right)_{k-1/k-1}} \\ \dfrac{q_{xv}}{\left(\sqrt{p_x\, p_v}\right)_{k-1/k-1}} & \dfrac{q_v}{\left(p_v\right)_{k-1/k-1}} \end{bmatrix} \quad . \quad \text{(C.16)}$$